



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 116827

TO: Fozia Hamud
Location: rem/4d64/4c70
Art Unit: 1647
Monday, March 15, 2004

Case Serial Number: 09/997573

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Hamud,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

STIC-Biotech/ChemLib

116829

From: Hamud, Fozia
Sent: Friday, March 12, 2004 4:30 PM
To: STIC-Biotech/ChemLib
Subject: sequence search for 09/997,573

Kindly search SEQ IID NO:418 of 09/997,573 against commercial and interference data abses. Thanks.

Fozia Hamud
Patent Examiner
Art Unit 1647
Remsen, Room 4D64
Mail Box Remsen-4C70
272-0884

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 3/15/04
Date Completed: 3/15/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 45p
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 15, 2004, 08:19:27 ; Search time 60 Seconds
(without alignments)
932.407 Million cell updates/sec

Title: US-09-997-573-418
Perfect score: 1031
Sequence: 1 MATLWGGLRLGLSLSCL.....WKLQVQRKSVFDRHVLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	100.0	198	3 AAY88275	Aay88275 Human TAN
2	1031	100.0	198	3 AAY66762	Aay66762 Membrane-
3	1031	100.0	198	3 AAY87231	Aay87231 Human sig
4	1031	100.0	198	3 AAY78807	Aay78807 Hydrophob
5	1031	100.0	198	4 AAM93740	Aam93740 Human pol
6	1031	100.0	198	4 AAB50966	Aab50966 Human pol
7	1031	100.0	198	4 AAB20120	Aab20120 Human imm
8	1031	100.0	198	4 AAM38735	Aam38735 Human pol
9	1031	100.0	198	4 AAB50926	Aab50926 Human PRO
10	1031	100.0	198	4 AAB565285	Aab565285 Human PRO
11	1031	100.0	198	5 ABP61428	Abp61428 Human NF-
12	1031	100.0	198	6 ABUS8100	Abu58100 Human PRO
13	1031	100.0	198	6 ABUS9178	Abu59178 Novel hum
14	1031	100.0	198	6 ABUS2690	Abu82690 Human sec
15	1031	100.0	198	6 ABUS60609	Abu60609 Human sec
16	1031	100.0	198	6 ABUI3991	Abui3991 Human PRO
17	1031	100.0	198	6 ABU72576	Abu72576 Novel hum
18	1031	100.0	198	6 ABU71432	Abu71432 Human neo
19	1031	100.0	198	6 ABUS9325	Abu59325 Human sec
20	1031	100.0	198	6 ABUS9325	Abu59325 Human PRO
21	1031	100.0	198	6 ABUS9031	Abu59031 Human sec
22	1031	100.0	198	6 ABUS2409	Abu32409 Novel hum
23	1031	100.0	198	6 ABUS9474	Abu59474 Novel hum
24	1031	100.0	198	6 ABUS2240	Abu52240 Novel hum
25	1031	100.0	198	6 ABUI0946	Abui0946 Human PRO

26	1031	100.0	198	6 ABUS1698	Abu81698 Novel hum
27	1031	100.0	198	6 ABUS8637	Abu8637 Human sec
28	1031	100.0	198	6 ABO34151	Abu34151 Human PRO
29	1031	100.0	198	6 ADA37929	Ada37929 Human sec
30	1031	100.0	198	6 ADA21615	Ada21615 Human sec
31	1031	100.0	198	6 ADA10402	Ada10402 Human sec
32	1031	100.0	198	6 ADA17946	Ada17946 Human PRO
33	1031	100.0	198	6 ADA28054	Ada28054 Human sec
34	1031	100.0	198	6 ADA94634	Ada94634 Human sec
35	1031	100.0	198	6 ADA38859	Ada38859 Human sec
36	1031	100.0	198	6 ADA92980	Ada92980 Human sec
37	1031	100.0	198	7 ABO53237	Abos53237 Human sec
38	1031	100.0	198	7 ADA22541	Ada22541 Human sec
39	1031	100.0	198	7 ABO22607	Abc22607 Human sec
40	1031	100.0	198	7 ADA06707	Ada06707 Human sec
41	1031	100.0	198	7 ADA39400	Ada39400 Human sec
42	1031	100.0	198	7 ADB96426	Adb96426 Human PRO
43	1031	100.0	198	7 ADCS7898	Acds7898 Human PRO
44	1031	100.0	198	7 ADCS5262	Acds5262 Human PRO
45	1031	100.0	198	7 ADC12129	Adc12129 Human sec

ALIGNMENTS

Sequence
'A'

RESULT 1
AAY88275
ID AAY88275 standard; protein; 198 AA.
XX
AC AAY88275;
XX
DT 16-OCT-2000 (first entry)
XX
DE Human TANGO 184 protein.
XX
KW TANGO 180; TANGO 181; TANGO 182; TANGO 183; TANGO 184; TANGO 185;
KW TANGO 186; TANGO 188; TANGO 189; TANGO 215; TANGO 187; human; murine;
KW secreted protein; transmembrane protein; gene therapy; vaccine;
KW diagnosis; treatment; detection.
XX
OS Homo sapiens.
XX
PN WQ200018904-A2.
XX
PD 06-APR-2000.
XX
PF 30-SEP-1999; 99WO-US022817.
XX
PR 30-SEP-1998; 98US-00164220.
XX
PA 02-OCT-1998; 98US-00164169.
XX
PI (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI Barnes TM;
XX
DR WPI; 2000-293144/25.
XX
DR N-PSDB; AAA39941, AAA39942.
XX

Isolated nucleic acids encoding TANGO polypeptides useful for preventing, diagnosing and treating diseases associated with inappropriate protein expression.
Claim 9; Fig 9; 249pp; English.
This invention describes novel human and murine nucleic acids encoding TANGO polypeptides (which are either wholly secreted or transmembrane proteins) which can be used for gene therapy and/or vaccination. The peptides are designated TANGO 180 to TANGO 189 and TANGO 215. The nucleic acids may be used to produce TANGO 180 to TANGO 189 and TANGO 215 polypeptides according to standard recombinant DNA methodologies. They may also be used to detect and quantify the presence of TANGO nucleic acids in a sample and therefore identify or diagnose diseases associated with inappropriate TANGO expression (e.g. diseases related to over or

See over

Sequence Comparison

CC under expression of the polypeptides or the expression of inactive
CC polypeptides). The nucleic acids and the polypeptides they encode may be
CC used according to standard gene therapy protocols, to treat diseases
CC associated with inappropriate TANGO expression by supplementing a
CC patient's own production of the polypeptide of to rectify mutations that
CC may result in expression of an abnormally active polypeptide. The
CC polypeptides may also be used to identify and produce agonists and
CC antagonists of TANGO expression and activity which may be used to
CC modulate TANGO related processes and diseases. The polypeptides are
CC particularly useful for use as antigens for producing antibodies to TANGO
CC proteins which may be used for inhibiting the activity of TANGO proteins.
CC They may also be used to detect and quantify the presence of TANGO
CC proteins in samples and therefore identify patients in whom the protein
CC is over- or under-expressed. This sequence represents the human TANGO 194
CC protein described in the method of the invention
XX
SQ

Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;

Best local Similarity 100.0%; Pred. No. 1.4e-104;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLRLGSLLSCLALSLLVLLAQLSDAAKNFEDVRCKICPPYKENSCHIYNKN 60

Db 1 MATLWGLRLGSLLSCLALSLLVLLAQLSDAAKNFEDVRCKICPPYKENSCHIYNKN 60

QY 61 ISOKDCDCLHVVPEVPGDPVAYCLRCCKYERSSTIKVTIIYLSILGLLLYMW 120

Db 61 ISOKDCDCLHVVPEVPGDPVAYCLRCCKYERSSTIKVTIIYLSILGLLLYMW 120

QY 121 YLTLVEPIKRLFGHAQLIQSDDDIGDHQPFANAHDLVARSBRANVLNKVEYAQRWK 180

Db 121 YLTLVEPIKRLFGHAQLIQSDDDIGDHQPFANAHDLVARSBRANVLNKVEYAQRWK 180

QY 181 LOVQQRKSVFDRHVLS 198

Db 181 LOVQQRKSVFDRHVLS 198

RESULT 2

AAV66762

ID AAV66762 standard; protein; 198 AA.

AC AAV66762;

XX 05-APR-2000 (first entry)

XX Membrane-bound protein PRO1375.

XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;

KW pharmaceutical; receptor immunoadhesin; gene mapping.

XX Homo sapiens.

XX WO9963088-A2.

XX 09-DEC-1999.

XX 02-JUN-1999; 99WO-US012252.

XX 02-JUN-1998; 98US-0087607P.

XX 02-JUN-1998; 98US-0087609P.

XX 03-JUN-1998; 98US-0087759P.

XX 04-JUN-1998; 98US-0088021P.

XX 04-JUN-1998; 98US-0088025P.

XX 04-JUN-1998; 98US-0088028P.

XX 04-JUN-1998; 98US-0088030P.

XX 04-JUN-1998; 98US-0088033P.

XX 05-JUN-1998; 98US-0088167P.

XX 05-JUN-1998; 98US-0088202P.

Sequence Comparison

PR 05-JUN-1998; 98US-0088212P.

PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-0088655P.

PR 10-JUN-1998; 98US-0088722P.

PR 10-JUN-1998; 98US-0088730P.

PR 10-JUN-1998; 98US-0088734P.

PR 10-JUN-1998; 98US-0088738P.

PR 10-JUN-1998; 98US-0088740P.

PR 10-JUN-1998; 98US-0088741P.

PR 10-JUN-1998; 98US-0088742P.

PR 10-JUN-1998; 98US-0088810P.

PR 10-JUN-1998; 98US-0088811P.

PR 10-JUN-1998; 98US-0088824P.

PR 10-JUN-1998; 98US-0088825P.

PR 10-JUN-1998; 98US-0088826P.

PR 11-JUN-1998; 98US-0088858P.

PR 11-JUN-1998; 98US-0088851P.

PR 11-JUN-1998; 98US-0088861P.

PR 11-JUN-1998; 98US-0088863P.

PR 11-JUN-1998; 98US-0088876P.

PR 12-JUN-1998; 98US-0089030P.

PR 12-JUN-1998; 98US-0089105P.

PR 16-JUN-1998; 98US-0089440P.

PR 16-JUN-1998; 98US-0089512P.

PR 16-JUN-1998; 98US-0089514P.

PR 17-JUN-1998; 98US-0089532P.

PR 17-JUN-1998; 98US-0089538P.

PR 17-JUN-1998; 98US-0089598P.

PR 17-JUN-1998; 98US-0089599P.

PR 17-JUN-1998; 98US-0089600P.

PR 17-JUN-1998; 98US-0089653P.

PR 18-JUN-1998; 98US-0089901P.

PR 18-JUN-1998; 98US-0089907P.

PR 18-JUN-1998; 98US-0089908P.

PR 19-JUN-1998; 98US-0089947P.

PR 19-JUN-1998; 98US-0089948P.

PR 19-JUN-1998; 98US-0089952P.

PR 22-JUN-1998; 98US-0090246P.

PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.

PR 23-JUN-1998; 98US-0090349P.

PR 23-JUN-1998; 98US-0090355P.

PR 24-JUN-1998; 98US-0090429P.

PR 24-JUN-1998; 98US-0090431P.

PR 24-JUN-1998; 98US-0090435P.

PR 24-JUN-1998; 98US-0090444P.

PR 24-JUN-1998; 98US-0090445P.

PR 24-JUN-1998; 98US-0090461P.

PR 24-JUN-1998; 98US-0090472P.

PR 24-JUN-1998; 98US-0090535P.

PR 24-JUN-1998; 98US-0090538P.

PR 24-JUN-1998; 98US-0090540P.

PR 24-JUN-1998; 98US-0090557P.

PR 25-JUN-1998; 98US-0090676P.

PR 25-JUN-1998; 98US-0090678P.

PR 25-JUN-1998; 98US-0090688P.

PR 25-JUN-1998; 98US-0090690P.

PR 25-JUN-1998; 98US-0090691P.

PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.

PR 25-JUN-1998; 98US-0090696P.

PR 26-JUN-1998; 98US-0090862P.

PR 26-JUN-1998; 98US-0090863P.

PR 01-JUL-1998; 98US-0091358P.

PR 01-JUL-1998; 98US-0091360P.

PR 02-JUL-1998; 98US-0091478P.

PR 02-JUL-1998; 98US-0091486P.

PR 02-JUL-1998; 98US-0091519P.

PR 02-JUL-1998; 98US-0091544P.

PR 02-JUL-1998; 98US-0091626P.

PR 02-JUL-1998; 98US-0091628P.

PR 02-JUL-1998; 98US-0091633P.

PR 02-JUL-1998; 98US-0091646P.

PR 02-JUL-1998; 98US-0091673P.

Sequence Comparison

07-JUL-1998; 98US-0091978P.
07-JUL-1998; 98US-0091982P.
09-JUL-1998; 98US-0092182P.
10-JUL-1998; 98US-0092472P.
20-JUL-1998; 98US-0093339P.
30-JUL-1998; 98US-0094651P.
04-AUG-1998; 98US-0095282P.
04-AUG-1998; 98US-0095285P.
04-AUG-1998; 98US-0095301P.
04-AUG-1998; 98US-0095302P.
04-AUG-1998; 98US-0095318P.
04-AUG-1998; 98US-0095321P.
04-AUG-1998; 98US-0095325P.
10-AUG-1998; 98US-0095916P.
10-AUG-1998; 98US-0095929P.
11-AUG-1998; 98US-0096012P.
11-AUG-1998; 98US-0096143P.
11-AUG-1998; 98US-0096146P.
17-AUG-1998; 98US-0096329P.
17-AUG-1998; 98US-0096757P.
17-AUG-1998; 98US-0096766P.
17-AUG-1998; 98US-0096768P.
17-AUG-1998; 98US-0096773P.
17-AUG-1998; 98US-0096791P.
17-AUG-1998; 98US-0096867P.
17-AUG-1998; 98US-0096891P.
17-AUG-1998; 98US-0096894P.
17-AUG-1998; 98US-0096895P.
17-AUG-1998; 98US-0096897P.
18-AUG-1998; 98US-0096949P.
18-AUG-1998; 98US-0096950P.
18-AUG-1998; 98US-0096959P.
18-AUG-1998; 98US-0096960P.
18-AUG-1998; 98US-0097022P.
19-AUG-1998; 98US-0097141P.
20-AUG-1998; 98US-0097218P.
24-AUG-1998; 98US-0097661P.
26-AUG-1998; 98US-0097951P.
26-AUG-1998; 98US-0097952P.
26-AUG-1998; 98US-0097955P.
26-AUG-1998; 98US-0097957P.
26-AUG-1998; 98US-0097974P.
26-AUG-1998; 98US-0097978P.
26-AUG-1998; 98US-0097979P.
26-AUG-1998; 98US-0097986P.
26-AUG-1998; 98US-0098014P.
31-AUG-1998; 98US-0098525P.
16-SEP-1998; 98US-0100634P.
12-JAN-1999; 99US-0115565P.

(GETH) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
Wood WI, Yuan J;

WPI; 2000-072883/06.
N-PSDB; AA265108.

Membrane-bound proteins and related nucleotide sequences.

Claim 12; Fig 300; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences

CC are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques

XX Sequence 198 AA;
SQ

Query Match 100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLRLGSLISLCLALSLLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNKN 60
DB 1 MATLWGLRLGSLISLCLALSLLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNKN 60
QY 61 ISQKDCDCLHVVEMVPRGPDVEAYCLRCECKYEERSVTIKVIIIVLSILGLLLYMV 120
DB 61 ISQKDCDCLHVVEMVPRGPDVEAYCLRCECKYEERSVTIKVIIIVLSILGLLLYMV 120
QY 121 YLTIVPEILKRLFGHAQLIQSDDDIGDHQPFANAHDVLAHSRANVKNVEYAQQRWK 180
DB 121 YLTIVPEILKRLFGHAQLIQSDDDIGDHQPFANAHDVLAHSRANVKNVEYAQQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198

Sequence Comparison

RESULT 3
AA87231
ID AA87231 standard; protein; 198 AA.

XX AA87231;

XX DT 11-MAY-2000 (first entry)

XX Human signal peptide containing protein HSPB-8 SEQ ID NO:8.

Human; signal peptide-containing protein; HSPB; diagnosis; cancer; inflammation; cardiovascular disease; anticancer; anti-inflammatory; antimicrobial; nontoxic; neuroprotective; cardiovascular; hepatotropic; antasthmatic; gene therapy; cell proliferation; neurological disorder; KW reproductive disorder; developmental disorder; arteriosclerosis; KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia; asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia; Parkinson's disease; Huntington's disease; ovulatory defect; muscular dystrophy.

XX Homo sapiens.

XX WC2000000610-A2.

XX 06-JAN-2000.

XX 25-JUN-1999; 99WO-US014484.

XX 26-JUN-1998; 98US-0090762P.

XX 31-JUL-1998; 98US-0094983P.

XX 01-OCT-1998; 98US-0102686P.

XX 11-DEC-1998; 98US-0112129P.

XX (INCY-) INCYTE PHARM INC.

XX Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI Bandman O;

XX WPI; 2000-160673/14.
XX N-PSDB; AA298116.

XX New human signal peptide-containing proteins useful in treatment, prevention and diagnosis of e.g. cancer, inflammation and cardiovascular disease.

See over

Sequence Comparison

```
XX AC AAM93740;
XX AC
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide, SEQ ID NO: 3711.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX OS Homo sapiens.
XX PN EPI130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX PA (HELI-) HELIX RES INST.
XX OTa T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX PR WPI; 2001-524255/58.
XX DR N-PSDB; RAK94692.
XX XX
XX PR 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX PS Claim 8; SEQ ID NO 3711; 1380pp + Sequence Listing; English.
XX CC The invention relates to primers for synthesizing full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesizing the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesised by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a polypeptide encoded by a full length
XX CC human cDNA of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in CD-ROM
XX CC format directly from EPO
XX SQ Sequence 198 AA;
XX
XX Query Match 100.0%; Score 1031; DB 4; Length 198;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-104;
XX Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGGLRLGSLLSCLALSULLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
DB 1 MATLWGGLRLGSLLSCLALSULLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
QY 61 ISQKDCDLHVVEPVRGPDVEAYCLRCCKEYERSSVTIKVTIIIVLSILGLLLYV 120
DB 61 ISQKDCDLHVVEPVRGPDVEAYCLRCCKEYERSSVTIKVTIIIVLSILGLLLYV 120
QY 121 YLTLPVPIKRLFGHAQLIQSDDDIGDHPFANAHVLRSRANVINKVEYAQRWK 180
DB 121 YLTLPVPIKRLFGHAQLIQSDDDIGDHPFANAHVLRSRANVINKVEYAQRWK 180
QY 181 LQVQEQKSVFDRHVLS 198
DB 181 LQVQEQKSVFDRHVLS 198
```

RESULT 6
AAB50966
ID AAB50966 standard; protein; 198 AA.

XX AAB50966;
XX AC
XX DT 21-MAR-2001 (first entry)
XX DE Human PRO1375 protein.
XX KW Human; PRO; cytostatic; neuroprotective; respiratory general;
XX KW antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
XX KW PRO agonist; cancer; inflammatory disorder; immunological disorder.
XX OS Homo sapiens.
XX PN WO200073348-A2.
XX PD 07-DEC-2000.
XX PF 30-MAY-2000; 2000WO-US014941.
XX PR 02-JUN-1999; 99WO-US012352.
XX PR 22-JUN-1999; 99US-0140650P.
XX PR 23-JUN-1999; 99US-0141037P.
XX PR 20-JUL-1999; 99US-0144758P.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 08-SEP-1999; 99WO-US020594.
XX PR 29-OCT-1999; 99US-0182508P.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 01-DEC-1999; 99WO-US028634.
XX PR 02-DEC-1999; 99US-0170262P.
XX PR 16-DEC-1999; 99WO-US030095.
XX PR 20-DEC-1999; 99WO-US030999.
XX PR 06-JAN-2000; 2000WO-US000376.
XX PR 11-FEB-2000; 2000WO-US003565.
XX PR 18-FEB-2000; 2000WO-US004341.
XX PR 02-MAR-2000; 2000WO-US005841.
XX PR 03-MAR-2000; 2000US-0187202P.
XX PR 10-MAR-2000; 2000WO-US006319.
XX PR 15-MAR-2000; 2000WO-US006884.
XX PR 30-MAR-2000; 2000WO-US008439.
XX PR 17-MAY-2000; 2000WO-US013705.
XX PA (GETH) GENENTECH INC.
XX PI Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;
XX PI Shelton DL, Smith V, Watanabe CK, Wood WI;
XX DR WPI; 2001-016509/02.
XX DR N-PSDB; AAC91568.
XX PT Twenty eight nucleic acids encoding PRO polypeptides which are useful for
XX PT treating various tumors, e.g. breast cancer, and other inflammatory,
XX PT angiogenic and immunological disorders.
XX PS Claim 31; Fig 32; 189pp; English.
XX CC The present sequence is one of twenty eight novel PRO polypeptides. The
XX CC PRO polypeptides and their agonists, including antibodies, peptides, and
XX CC small molecule agonists, may be used to treat various tumors, e.g.,
XX CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
XX CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
XX CC central nervous system cancer, melanoma or leukaemia. They are also
XX CC useful for treating other disorders such as neuronal, glial, astrocytal,
XX CC hypothalamic and other glandular, macrophagal, epithelial, stromal and
XX CC blastocoealic disorders, and inflammatory, angiogenic and immunological
XX CC disorders
XX SQ Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
DB 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
QY 61 ISQKDCDCLHVPMPVGRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
DB 61 ISQKDCDCLHVPMPVGRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
QY 121 YLTLPVLPKRLFGHAQLIQSDDDDIGDHPFANAHDVLAARSRLANVKNVEYAQRWK 180
DB 121 YLTLPVLPKRLFGHAQLIQSDDDDIGDHPFANAHDVLAARSRLANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 7
ID AAB20120
ID AAB20120 standard; protein; 198 AA.
XX
AC AAB20120;
XX
DT 30-APR-2001 (first entry)
XX
DE Human immunostimulant PRO1375.
XX
KW PRO1375; UNQ712; human; immune disease; autoimmune disease;
KW antineumatic; antiarthritic; antiinflammatory; antianemic;
KW immunosuppressive; antithyroid; antidiabetic; neuroprotective;
KW hepatotropic; virucide; dermatological; antipsoriatic; antiasthmatic;
KW antiallergic; immunostimulant.
XX
OS Homo sapiens.
XX
FH Key
FT Domain
FT 11..28 Location/Qualifiers
FT /note= "transmembrane domain type II"
FT Modified-site 12..18
FT /note= "N-myristoylation site"
FT Modified-site 60..64
FT /note= "Asn is N-glycosylated"
FT Modified-site 78..86
FT /note= "tyrosine kinase phosphorylation site"
FT Domain 103..125
FT /note= "transmembrane domain"
XX
PN WO200105972-A1.
XX
PD 25-JAN-2001.
XX
PF 15-MAR-2000; 2000WO-US006884.
XX
PR 20-JUL-1999; 99US-0144758P.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Fong S, Goddard A, Godowski PJ;
PI Gurney AL, Hillan KJ, Mark MR, Marsters SA, Pitti RM, Tumas D;
PI Watanabe CK, Wood WI;
XX
XX WPI; 2001-103149/11.
DR N-PSDB; AAF30062.
XX
XX New PRO polypeptides, nucleic acids and (ant)agonists, useful for
PT diagnosing and treating immune-related disorders, such as multiple
PT sclerosis, rheumatoid arthritis and diabetes.
XX
PS Claim 20; Fig 26; 127pp; English.
XX
CC The present sequence is that of novel human immunomodulator PRO1375
CC (UNQ712), as deduced from cDNA (see AAF30062) isolated from a human

CC pancreas tissue library. PRO1375 has a mol.wt. of 22 kDa and a pI of
CC 8.47. The invention provides polynucleotides (see AAF30050-62) encoding
CC novel human PRO proteins (see AAB20108-20) including PRO1375. Claimed
CC compositions comprising these proteins or their agonists are useful for
CC increasing infiltration of inflammatory cells into a tissue of a mammal,
CC stimulating or enhancing an immune response, or increasing the
CC proliferation of T-lymphocytes in a mammal in response to an antigen.
CC Claimed compositions comprising a PRO polypeptide or its antagonist have
CC the opposite effect. A claimed method for treating an immune related
CC disorder, such as a T cell disorder, involves administering a PRO
CC polypeptide, an agonist antibody or an antagonist antibody. The disorder
CC is selected from systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis, systemic
CC sclerosis, idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, demyelinated diseases (such as multiple sclerosis), autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, sclerosing cholangitis, inflammatory bowel disease (ulcerative
CC colitis and Crohn's disease), gluten-sensitive enteropathy, Whipple's
CC disease, (auto)immune-mediated skin diseases (such as bullous skin
CC disease, erythema multiforme and psoriasis), allergic diseases (such as
CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity and
CC urticaria), immunologic diseases of the lung and transplantation
CC associated diseases (such as graft rejection and graft-versus-host
CC disease) (all claimed). Claimed methods of diagnosing these disorders
CC comprise detecting the level of expression of the PRO gene. Also claimed
CC are a method of identifying a compound capable of inhibiting the
CC expression or activity of the PRO polypeptide, vectors, host cells,
CC antibodies, and a method of stimulating the proliferation of T
CC lymphocytes using PRO1375
SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
DB 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
QY 61 ISQKDCDCLHVPMPVGRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
DB 61 ISQKDCDCLHVPMPVGRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
QY 121 YLTLPVLPKRLFGHAQLIQSDDDDIGDHPFANAHDVLAARSRLANVKNVEYAQRWK 180
DB 121 YLTLPVLPKRLFGHAQLIQSDDDDIGDHPFANAHDVLAARSRLANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 8
AAB38735
ID AAB38735 standard; protein; 198 AA.
XX
AC AAB38735;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1880.
XX
XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:19:27 ; Search time 60 Seconds
(without alignments)
932.407 Million cell updates/sec

Title: US-09-997-573-418
Perfect score: 1031
Sequence: 1 MARLMGGLRLGLSLLSLCL.....WKLQVQRKSVFDRHVLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	100.0	198	3	AY88275 Human TAN
2	1031	100.0	198	3	AY66762 Membrane-
3	1031	100.0	198	3	AY87231 Human sig
4	1031	100.0	198	3	AY78807 Hydrophob
5	1031	100.0	198	4	AA93740 Human pol
6	1031	100.0	198	4	AB50966 Human PRO
7	1031	100.0	198	4	AB20120 Human imm
8	1031	100.0	198	4	AB38735 Human pol
9	1031	100.0	198	4	AB50926 Human PRO
10	1031	100.0	198	4	AB65285 Human PRO
11	1031	100.0	198	5	ABP61428 Human NF-
12	1031	100.0	198	6	ABU58100 Human PRO
13	1031	100.0	198	6	ABU59178 Novel hum
14	1031	100.0	198	6	ABU82690 Human sec
15	1031	100.0	198	6	ABU60609 Human sec
16	1031	100.0	198	6	ABU13991 Human PRO
17	1031	100.0	198	6	ABU72576 Novel hum
18	1031	100.0	198	6	ABU71432 Human neo
19	1031	100.0	198	6	ABU59325 Human sec
20	1031	100.0	198	6	ABO26022 Human PRO
21	1031	100.0	198	6	ABU59031 Human sec
22	1031	100.0	198	6	ABU92409 Novel hum
23	1031	100.0	198	6	ABU5474 Novel hum
24	1031	100.0	198	6	ABU92240 Novel hum
25	1031	100.0	198	6	ABU10946 Human PRO

26 1031 100.0 198 6 ABU81698 Novel hum
27 1031 100.0 198 6 ABU88637 Human sec
28 1031 100.0 198 6 ABO34151 Human PRO
29 1031 100.0 198 6 ADA37929 Human sec
30 1031 100.0 198 6 ADA21615 Human sec
31 1031 100.0 198 6 ADA10402 Human sec
32 1031 100.0 198 6 ADA17946 Human PRO
33 1031 100.0 198 6 ADA28054 Human sec
34 1031 100.0 198 6 ADA94634 Human sec
35 1031 100.0 198 6 ADA38859 Human sec
36 1031 100.0 198 6 ADA92980 Human sec
37 1031 100.0 198 7 ABO53237 Human sec
38 1031 100.0 198 7 ABO22541 Human sec
39 1031 100.0 198 7 ABO22607 Human sec
40 1031 100.0 198 7 ABO22607 Human sec
41 1031 100.0 198 7 ADA39400 Human sec
42 1031 100.0 198 7 ADB96426 Human PRO
43 1031 100.0 198 7 ADC57898 Human PRO
44 1031 100.0 198 7 ADC55262 Human PRO
45 1031 100.0 198 7 ADC12129 Human sec

ALIGNMENTS

RESULT 1
AAV88275
ID AAV88275 standard; protein; 198 AA.

AC AAV88275;

DT 16-OCT-2000 (first entry)

DE Human TANGO 184 protein.

XX TANGO 180; TANGO 181; TANGO 182; TANGO 183; TANGO 184; TANGO 185;

KW TANGO 186; TANGO 188; TANGO 189; TANGO 215; TANGO 187; human; murine;

KW secreted protein; transmembrane protein; gene therapy; vaccine;

KW diagnosis; treatment; detection.

OS Homo sapiens.

PN WO200018904-A2.

XX 06-APR-2000.

XX 30-SEP-1999; 99WO-US022817.

XX 30-SEP-1998; 98US-00164220.

XX 02-OCT-1998; 98US-00164169.

(MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

PI Barnes TM;

DR WPI: 2000-293144/25.

DR N-PSDB; AAA39941, AAA39942.

XX Isolated nucleic acids encoding TANGO polypeptides useful for preventing,

XX diagnosing and treating diseases associated with inappropriate protein

XX expression.

PS Claim 9; Fig 9; 249pp; English.

XX This invention describes novel human and murine nucleic acids encoding

XX TANGO polypeptides (which are either wholly secreted or transmembrane

XX proteins) which can be used for gene therapy and/or vaccination. The

XX peptides are designated TANGO 180 to TANGO 189 and TANGO 215. The nucleic

XX acids may be used to produce TANGO 180 to TANGO 189 and TANGO 215

XX polypeptides according to standard recombinant DNA methodologies. They

XX may also be used to detect and quantify the presence of TANGO nucleic

XX acids in a sample and therefore identify or diagnose diseases associated

XX with inappropriate TANGO expression (e.g. diseases related to over or

CC under expression of the polypeptides or the expression of inactive
CC polypeptides). The nucleic acids and the polypeptides they encode may be
CC used according to standard gene therapy protocols, to treat diseases
CC associated with inappropriate TANGO expression by supplementing a
CC patient's own production of the polypeptide of to rectify mutations that
CC may result in expression of an abnormally active polypeptide. The
CC polypeptides may also be used to identify and produce agonists and
CC antagonists of TANGO expression and activity which may be used to
CC modulate TANGO related processes and diseases. The polypeptides are
CC particularly useful for use as antigens for producing antibodies to TANGO
CC proteins which may be used for inhibiting the activity of TANGO proteins.
CC They may also be used to detect and quantify the presence of TANGO
CC proteins in samples and therefore identify patients in whom the protein
CC is over- or under-expressed. This sequence represents the human TANGO 184
CC protein described in the method of the invention

XX Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLLRIGLSLSCLSLVLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
DB 1 MATLWGGLLRIGLSLSCLSLVLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60

QY 61 ISQKDCDCLHVPMPVPRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYMV 120
DB 61 ISQKDCDCLHVPMPVPRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYMV 120

QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHVLAHSRANVINKVEYAQRWK 180
DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHVLAHSRANVINKVEYAQRWK 180

QY 181 LQVQEQKSVFDRHVLS 198
DB 181 LQVQEQKSVFDRHVLS 198

RESULT 2

ID AAY66762 standard; protein; 198 AA.

XX AAY66762;

XX 05-APR-2000 (first entry)

XX Membrane-bound protein PRO1375.

XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
XX pharmaceutical; receptor immunoadhesin; gene mapping.

XX Homo sapiens.

XX WO9963088-A2.

XX 09-DEC-1999

XX 02-JUN-1999; 99WO-US012252.

XX 02-JUN-1998; 98US-0087607P.

XX 02-JUN-1998; 98US-0087609P.

XX 02-JUN-1998; 98US-0087759P.

XX 03-JUN-1998; 98US-0087827P.

XX 04-JUN-1998; 98US-0088021P.

XX 04-JUN-1998; 98US-0088025P.

XX 04-JUN-1998; 98US-0088028P.

XX 04-JUN-1998; 98US-0088029P.

XX 04-JUN-1998; 98US-0088030P.

XX 04-JUN-1998; 98US-0088033P.

XX 05-JUN-1998; 98US-0088326P.

XX 05-JUN-1998; 98US-0088167P.

XX 05-JUN-1998; 98US-0088202P.

PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088730P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088741P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088858P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 12-JUN-1998; 98US-0089440P.
PR 12-JUN-1998; 98US-0089512P.
PR 12-JUN-1998; 98US-0089514P.
PR 12-JUN-1998; 98US-0089532P.
PR 12-JUN-1998; 98US-0089538P.
PR 12-JUN-1998; 98US-0089598P.
PR 12-JUN-1998; 98US-0089599P.
PR 12-JUN-1998; 98US-0089600P.
PR 12-JUN-1998; 98US-0089653P.
PR 12-JUN-1998; 98US-0089801P.
PR 12-JUN-1998; 98US-0089907P.
PR 12-JUN-1998; 98US-0089908P.
PR 12-JUN-1998; 98US-0089947P.
PR 12-JUN-1998; 98US-0089948P.
PR 12-JUN-1998; 98US-0089952P.
PR 12-JUN-1998; 98US-0090246P.
PR 12-JUN-1998; 98US-0090252P.
PR 12-JUN-1998; 98US-0090254P.
PR 12-JUN-1998; 98US-0090349P.
PR 12-JUN-1998; 98US-0090355P.
PR 12-JUN-1998; 98US-0090429P.
PR 12-JUN-1998; 98US-0090431P.
PR 12-JUN-1998; 98US-0090435P.
PR 12-JUN-1998; 98US-0090444P.
PR 12-JUN-1998; 98US-0090445P.
PR 12-JUN-1998; 98US-0090461P.
PR 12-JUN-1998; 98US-0090472P.
PR 12-JUN-1998; 98US-0090535P.
PR 12-JUN-1998; 98US-0090538P.
PR 12-JUN-1998; 98US-0090540P.
PR 12-JUN-1998; 98US-0090557P.
PR 12-JUN-1998; 98US-0090676P.
PR 12-JUN-1998; 98US-0090678P.
PR 12-JUN-1998; 98US-0090688P.
PR 12-JUN-1998; 98US-0090690P.
PR 12-JUN-1998; 98US-0090691P.
PR 12-JUN-1998; 98US-0090694P.
PR 12-JUN-1998; 98US-0090695P.
PR 12-JUN-1998; 98US-0090696P.
PR 12-JUN-1998; 98US-0090862P.
PR 12-JUN-1998; 98US-0090863P.
PR 12-JUN-1998; 98US-0091358P.
PR 12-JUN-1998; 98US-0091360P.
PR 12-JUN-1998; 98US-0091478P.
PR 12-JUN-1998; 98US-0091486P.
PR 12-JUN-1998; 98US-0091519P.
PR 12-JUN-1998; 98US-0091544P.
PR 12-JUN-1998; 98US-0091626P.
PR 12-JUN-1998; 98US-0091628P.
PR 12-JUN-1998; 98US-0091633P.
PR 12-JUN-1998; 98US-0091646P.
PR 12-JUN-1998; 98US-0091673P.

are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques

Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLLSCLALS VLLLAQLSDAAKNFEDVRCKICPPKYKNSGHYYNKN 60
DB 1 MATLWGGLRLGSLLSCLALS VLLLAQLSDAAKNFEDVRCKICPPKYKNSGHYYNKN 60
QY 61 ISQXDCDCLHVPEMPVGRGPDEAYCRLRCECKYEERSVTIKVIIYLSILGLLLLYMV 120
DB 61 ISQXDCDCLHVPEMPVGRGPDEAYCRLRCECKYEERSVTIKVIIYLSILGLLLLYMV 120
QY 121 YLTTLVEPILKERLFGHAQLIQSDDDIGDHQPFANAHVDVLARSRRANVNKVEYAQQRWK 180
DB 121 YLTTLVEPILKERLFGHAQLIQSDDDIGDHQPFANAHVDVLARSRRANVNKVEYAQQRWK 180
QY 181 LQVQEQRKSVPDRHWLS 198
DB 181 LQVQEQRKSVPDRHWLS 198

RESULT 3
AAY87231
ID AAY87231 standard; protein; 198 AA.
AC AAY87231;
XX
DT 11-MAY-2000 (first entry)
DE Human signal peptide containing protein HSPB-8 SEQ ID NO:8.
KW Human; signal peptide-containing protein; HSPB; diagnosis; cancer;
inflammation; cardiovascular disease; anticancer; anti-inflammatory;
antimicrobial; nontoxic; neuroprotective; cardiovascular; hepatotropic;
antiasthmatic; gene therapy; cell proliferation; neurological disorder;
reproductive disorder; developmental disorder; arteriosclerosis;
cirrhosis; psoriasis; acquired immune deficiency syndrome; anemia;
asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
Parkinson's disease; Huntington's diseases; ovulatory defect;
muscular dystrophy.
XX
OS Homo sapiens.
XX
PN WO200000610-A2.
PD
XX 06-JAN-2000.
XX
PF 25-JUN-1999; 99WO-US014484.
XX
PR 26-JUN-1998; 98US-0090762P.
PR 31-JUL-1998; 98US-0094983P.
PR 01-OCT-1998; 98US-0102686P.
PR 11-DEC-1998; 98US-0112129P.
XX
PA (INCY) INCYTE PHARM INC.
XX
PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn KR;
PI Akerbloom IL, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI Bandman O;
XX
DR WPI: 2000-160673/14.
DR N-PSDB; AA298116.
PT New human signal peptide-containing proteins useful in treatment,
PT prevention and diagnosis of e.g. cancer, inflammation and cardiovascular
PT disease.

XX PS Claim 1; Page 165; 327pp; English.

XX CC AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the

XX CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have

XX CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,

XX CC neuroprotective, cardiovascular and antiasthmatic activities, and can be

XX CC used in gene therapy. HSPs can be used to treat or prevent disorders

XX CC associated with decreased activity or function of HSP. Antagonists of

XX CC HSP are used to treat or prevent disorders associated with increased

XX CC activity or function of HSP. Such diseases include cell proliferation

XX CC (including cancer), inflammation, cardiovascular, neurological,

XX CC reproductive or developmental disorders, (e.g. arteriosclerosis,

XX CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,

XX CC asthma, Crohn's disease, microbial or other infections, congestive or

XX CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's

XX CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP

XX CC nucleic acids can be used for the recombinant production of HSP, for

XX CC detecting HSP in standard hybridisation and amplification assays (for

XX CC diagnosis and monitoring), in gene therapy, as antisense, triplex-forming

XX CC or ribozyme therapeutics, for detecting related sequences or genetic

XX CC variations, and for chromosomal mapping. HSP are also used to raise

XX CC specific antibodies (Ab) and to screen for agonists and antagonists

XX CC (potential therapeutic agents). Ab are used to diagnose, or monitor, HSP

XX CC -related diseases (in usual immunoassays), as therapeutic antagonists, in

XX CC competitive drug screens, and for purification of HSP from natural

XX CC sources

XX SQ Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;

Best Local Similarity 100.0%; Pred. No. 1.4e-104;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60

DB 1 MATLWGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60

QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSVTKVTIIYLSILGLLLYV 120

DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSVTKVTIIYLSILGLLLYV 120

QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHOPFANAHDVLARSRSRANVLNKEVYAOQRWK 180

DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHOPFANAHDVLARSRSRANVLNKEVYAOQRWK 180

QY 181 LQVQQRKSVFDRHVLS 198

DB 181 LQVQQRKSVFDRHVLS 198

RESULT 4

AAV78807

ID AAY78807 standard; protein; 198 AA.

XX AC AAY78807;

XX DT 09-MAY-2000 (first entry)

XX DE Hydrophobic domain containing protein clone HP10529 protein sequence.

XX KW Hydrophobic domain; clone HP10529; nutritional supplement; SCID; HIV;

XX KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;

XX KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;

XX KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;

XX KW anti-inflammatory; infection; bodily characteristic.

XX OS Homo sapiens.

XX PN WO200000506-A2.

XX PD 06-JAN-2000.

PF 18-JUN-1999; 99WO-JP003242.

XX PR 26-JUN-1998; 98JP-00180008.

XX PA (SAGA) SAGAMI CHEM RES CENT.

XX PA (PROT-) PROTEGENE INC.

XX PI Kato S, Kimura T;

XX DR WPI; 2000-160665/14.

DR N-PSDB; AAZ90044, AAZ90054.

XX Novel human proteins having hydrophobic domains used for research and

PT diagnostic purposes.

XX Claim 1; Page 79-80; 117pp; English.

XX This sequence represents the hydrophobic domain containing protein, clone

CC HP10529 protein sequence. The sequence is isolated from a human

CC osteosarcoma cell line Soas-2. The invention relates to human proteins

CC with hydrophobic domains, the DNA and the cDNA encoding them. The

CC polynucleotides and proteins are predicted to have biological activities

CC which make them suitable for treating, preventing or ameliorating medical

CC conditions in humans and animals. Suggested activities include

CC nutritional activity (nutritional source or supplement); cytokine and

CC cell proliferation/differentiation activity; immune stimulating (e.g. as

CC vaccines) or suppressing activity (e.g. to treat various immune

CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic

CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary

CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin

CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease

CC and autoimmune inflammatory eye disease, as well as asthma, allergies and

CC organ transplantation; haematopoiesis regulating activity (e.g. in

CC treatment of myeloid or lymphoid cell deficiencies); tissue growth

CC activity (e.g. wound healing and tissue repair, ulcers, burns,

CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic

CC activity; haemostatic and thrombolytic activity (e.g. treating

CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and

CC tumour inhibition activity. The polynucleotides are also stated to be

CC useful for gene therapy. Other activities include inhibiting infections

CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,

CC malaria); effecting bodily characteristics such as, e.g. weight, colour,

CC skin, effecting biorhythms or cardiac cycles, enhancing fertility;

CC treatment of depression; treatment of pain; hormonal or endocrine

CC activity. The polynucleotides may also be used for recombinant expression

CC of the protein

XX SQ Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;

Best Local Similarity 100.0%; Pred. No. 1.4e-104;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60

DB 1 MATLWGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60

QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSVTKVTIIYLSILGLLLYV 120

DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSVTKVTIIYLSILGLLLYV 120

QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHOPFANAHDVLARSRSRANVLNKEVYAOQRWK 180

DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHOPFANAHDVLARSRSRANVLNKEVYAOQRWK 180

QY 181 LQVQQRKSVFDRHVLS 198

DB 181 LQVQQRKSVFDRHVLS 198

RESULT 5

AAV93740

ID AAM93740 standard; protein; 198 AA.

QY 1 MATWGGLRLGSLLSCLALSLLAQLSDAKNFEDVRCICPPYKENSCHYNNK 60
DB 1 MATWGGLRLGSLLSCLALSLLAQLSDAKNFEDVRCICPPYKENSCHYNNK 60
QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSANVKNVEYAQRWK 180
DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 7
ID AAB20120
AC AAB20120 standard; protein; 198 AA.
XX AAB20120;
DT 30-APR-2001 (first entry)
DE Human immunostimulant PRO1375.
XX PRO1375; UNQ712; human; immune disease; autoimmune disease;
KW antirheumatic; antiarthritic; antiinflammatory; antianaemic;
KW immunosuppressive; antithyroid; antidiabetic; neuroprotective;
KW hepatotropic; virucide; dermatological; antipsoriatic; antiasthmatic;
KW antiallergic; immunostimulant.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Domain 11..28
FT /note= "transmembrane domain type II"
FT Modified-site 12..18
FT /note= "N-myristoylation site"
FT Modified-site 60..64
FT /note= "Asn is N-glycosylated"
FT Modified-site 78..86
FT /note= "tyrosine kinase phosphorylation site"
FT Domain 103..125
FT /note= "transmembrane domain"
XX WO200105972-A1.
XX 25-JAN-2001.
XX 15-MAR-2000; 2000WO-US006884.
XX 20-JUL-1999; 99US-0144758P.
XX (GETH) GENENTECH INC.
XX Ashkenazi AJ, Baker KP, Fong S, Goddard A, Godowski PJ;
PI Gurney AL, Hillan KJ, Mark MR, Marsters SA, Pitti RM, Tumas D;
PI Watanabe CK, Wood WI;
XX WPI; 2001-103149/11.
DR N-PSDB; AAF30062.
XX New PRO polypeptides, nucleic acids and (ant)agonists, useful for
PT diagnosing and treating immune-related disorders, such as multiple
PT sclerosis, rheumatoid arthritis and diabetes.
XX Claim 20; Fig 26; 127pp; English.
XX The present sequence is that of novel human immunomodulator PRO1375
CC (UNQ712), as deduced from cDNA (see AAF30062) isolated from a human

CC pancreas tissue library. PRO1375 has a mol.wt. of 22 kDa and a pI of
CC 8.47. The invention provides polynucleotides (see AAF30050-62) encoding
CC novel human PRO proteins (see AAB20108-20) including PRO1375. Claimed
CC compositions comprising these proteins or their agonists are useful for
CC increasing infiltration of inflammatory cells into a tissue of a mammal,
CC stimulating or enhancing an immune response, or increasing the
CC proliferation of T-lymphocytes in a mammal in response to an antigen.
CC Claimed compositions comprising a PRO polypeptide or its antagonist have
CC the opposite effect. A claimed method for treating an immune related
CC disorder, such as a T cell disorder, involves administering a PRO
CC polypeptide, an agonist antibody or an antagonist antibody. The disorder
CC is selected from systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis, systemic
CC sclerosis, idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune renal
CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, demyelinated diseases (such as multiple sclerosis), autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, sclerosing cholangitis, inflammatory bowel disease (ulcerative
CC colitis and Crohn's disease), gluten-sensitive enteropathy, Whipple's
CC disease, (auto)immune-mediated skin diseases (such as bullous skin
CC disease, erythema multiforme and psoriasis), allergic diseases (such as
CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity and
CC urticaria), immunologic diseases of the lung and transplantation
CC associated diseases (such as graft rejection and graft-versus-host
CC disease) (all claimed). Claimed methods of diagnosing these disorders
CC comprise detecting the level of expression of the PRO gene. Also claimed
CC are a method of identifying a compound capable of inhibiting the
CC expression or activity of the PRO polypeptide, vectors, host cells,
CC antibodies, and a method of stimulating the proliferation of T
CC lymphocytes using PRO1375
XX Sequence 198 AA;
SQ
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATWGGLRLGSLLSCLALSLLAQLSDAKNFEDVRCICPPYKENSCHYNNK 60
DB 1 MATWGGLRLGSLLSCLALSLLAQLSDAKNFEDVRCICPPYKENSCHYNNK 60
QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYV 120
QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSANVKNVEYAQRWK 180
DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 8
ID AAM38735
AC AAM38735 standard; protein; 198 AA.
XX AAM38735;
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 1880.
XX Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
OS

XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-0052317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Auandi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Qa;
PI Zhou P, Goodrich R, Drmanac RT;
XX DR WPI; 2001-442253/47.
XX DR N-PSDB; AAI57891.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX PS Example 3; SEQ ID NO 1880; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104; Mismatches 0; Indels 0; Gaps 0;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGLRLGSLSLALSLLVLLLAQLSDAKNFEDVRCKICPPYKENSIGHIYNKN 60
Db 1 MATLWGLRLGSLSLALSLLVLLLAQLSDAKNFEDVRCKICPPYKENSIGHIYNKN 60
QY 61 ISQKDCDCLHVVPEPVPVGPVVEAYCLRCCKEYERSVTKVIIIVLSILGLLLYV 120
Db 61 ISQKDCDCLHVVPEPVPVGPVVEAYCLRCCKEYERSVTKVIIIVLSILGLLLYV 120
QY 121 YTLVPEPLKRLFGHQLIQSDDDIGDHQHPANAHDLVARSGRANVLNKVEYAQRWK 180
Db 121 YTLVPEPLKRLFGHQLIQSDDDIGDHQHPANAHDLVARSGRANVLNKVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
Db 181 LQVQQRKSVFDRHVLS 198
RESULT 9
AAB50926
ID AAB50926 standard; protein; 198 AA.

XX AC AAB50926;
XX DT 21-MAR-2001 (first entry)
XX DE Human PRO1375 protein.
XX KW Human; PRO; antiinflammatory; dermatological; antiarthritic;
KW antirheumatic; cardiac; antianemic; immunosuppressive; antithyroid;
KW antidiabetic; neuroprotective; hepatotropic; virucide;
KW antiallergic; antiasthmatic; immune related disorder;
KW hepatobiliary disease; autoimmune disease; allergy.
XX OS Homo sapiens.
XX PN WO200073452-A2.
XX PD 07-DEC-2000.
XX PF 02-JUN-2000; 2000WO-US015264.
XX PR 02-JUN-1999; 99WO-US012252.
XX PR 20-JUL-1999; 99US-0144732P.
XX PR 20-JUL-1999; 99US-0144758P.
XX PR 28-JUL-1999; 99US-0146222P.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 15-SEP-1999; 99WO-US021090.
XX PR 15-SEP-1999; 99WO-US021547.
XX PR 29-OCT-1999; 99US-0162506P.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 01-DEC-1999; 99WO-US028634.
XX PR 02-DEC-1999; 99WO-US028551.
XX PR 02-DEC-1999; 99WO-US028565.
XX PR 09-DEC-1999; 99US-0170262P.
XX PR 20-DEC-1999; 99WO-US030911.
XX PR 05-JAN-2000; 2000WO-US000219.
XX PR 11-FEB-2000; 2000WO-US000376.
XX PR 18-FEB-2000; 2000WO-US004341.
XX PR 22-FEB-2000; 2000WO-US004342.
XX PR 24-FEB-2000; 2000WO-US004914.
XX PR 24-FEB-2000; 2000WO-US005004.
XX PR 01-MAR-2000; 2000WO-US005601.
XX PR 02-MAR-2000; 2000WO-US005841.
XX PR 03-MAR-2000; 2000US-0187202P.
XX PR 15-MAR-2000; 2000WO-US006884.
XX PR 20-MAR-2000; 2000WO-US007377.
XX PR 21-MAR-2000; 2000WO-US007532.
XX PR 30-MAR-2000; 2000WO-US008439.
XX PR 17-MAY-2000; 2000WO-US013705.
XX PR 22-MAY-2000; 2000WO-US014042.
XX (GETH) GENENTECH INC.
XX PA Ashkenazi AJ, Baker KP, Chan B, Goddard A, Godowski PJ;
XX Gurney AL, Hebert C, Henzel W, Kabakoff RC, Shelton DL, Tumas D;
XX Watanabe CK, Wood WI;
XX WPI; 2001-025253/03.
XX N-PSDB; AAC91485.
XX PT Thirty three nucleic acids encoding PRO polypeptides which are useful in
XX the diagnosis and treatment of immune related disorders, e.g. systemic
XX lupus erythematosus, rheumatoid arthritis, osteoarthritis, thyroiditis
XX and diabetes mellitus.
XX PS Claim 58; Fig 50; 218pp; English.
XX CC The present sequence is one of thirty three novel PRO polypeptides. The
XX PRO polypeptides, anti-PRO antibodies, agonists and antagonists are
XX useful for treating and diagnosing immune related disorders such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX CC

CC juvenile chronic arthritis, spondyloarthropathies, systemic sclerosis,
 CC idiopathic inflammatory myopathies, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, demyelinating diseases of the central and peripheral nervous
 CC systems (such as multiple sclerosis, idiopathic demyelinating
 CC polyneuropathy or Guillain-Barre syndrome, and chronic inflammatory
 CC demyelinating polyneuropathy), hepatobiliary diseases (such as
 CC infectious, granulomatous chronic active hepatitis, primary biliary
 CC cirrhosis, autoimmune chronic active hepatitis and sclerosing cholangitis),
 CC inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's
 CC disease, autoimmune or immune-mediated skin diseases (such as bullous
 CC skin diseases, erythema multiforme, contact dermatitis, psoriasis),
 CC allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
 CC food hypersensitivity and urticaria), immunological diseases of the lung
 CC (such as eosinophilic pneumonias, idiopathic pulmonary fibrosis and
 CC hypersensitivity pneumonitis), transplantation associated diseases
 CC including graft rejection and graft-versus-host diseases
 XX
 XX Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 4; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.4e-104;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAKNFEDVRCKICPPYKENSCHIYKN 60
 DB 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAKNFEDVRCKICPPYKENSCHIYKN 60
 QY 61 ISQKDCDCLHVPEPMPVPGDVEAYCLRCCKYEERSSVTIKVTIIIVLSILGLLLYV 120
 DB 61 ISQKDCDCLHVPEPMPVPGDVEAYCLRCCKYEERSSVTIKVTIIIVLSILGLLLYV 120
 QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 QY 181 LQVQEQKSVFDRHVLS 198
 DB 181 LQVQEQKSVFDRHVLS 198

RESULT 10
 AAB65285
 ID AAB65285 standard; protein; 198 AA.
 AC AAB65285;

XX
 XX 02-APR-2001 (first entry)

XX Human PRO1375 (UNO712) protein sequence SEQ ID NO:418.

XX Human; secreted and transmembrane protein; PRO; cytostatic; cell death;
 KW cancer; chromosomal mapping; Gene mapping; tissue typing;
 KW diagnostic assay.

XX Homo sapiens.

XX WO2000073454-A1.

XX 07-DEC-2000.

XX 30-MAR-2000; 2000WO-US008439.

XX 02-JUN-1999; 99WO-US012252.
 PR 23-JUN-1999; 99US-0141037P.
 PR 07-JUL-1999; 99US-0143048P.
 PR 20-JUL-1999; 99US-0144758P.
 PR 26-JUL-1999; 99US-0145698P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 17-AUG-1999; 99US-0149336P.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.

PR 08-OCT-1999; 99US-0158663P.
 PR 30-NOV-1999; 99WO-US028313.
 PR 01-DEC-1999; 99WO-US028301.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 28-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.

(GETH) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Fong S, Gerber H, Gerritsen MP, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
 PI Zhang Z;

WPI; 2001-032150/04.

N-PSDB; AAF44254.

PR PRO polynucleotides used to produce polypeptides used to target bioactive
 PT molecules such as toxins, radiolabels or antibodies, to specific cells,
 PT to cause targeted cell death.

XX Claim 12; Fig 300; 935pp; English.

XX The present invention describes human secreted and transmembrane PRO
 CC proteins. The PRO proteins have cytostatic activity. The PRO proteins can
 CC be used for targeted delivery of bioactive molecules, such as toxins,
 CC radiolabels or antibodies, that cause cell death. PRO nucleotide
 CC sequences, and their fragments, can be used as hybridisation probes, in
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA and
 CC DNA. They may also be used to produce transgenic animals which are used
 CC to develop and screen therapeutically useful reagents. The PRO nucleotide
 CC and protein sequence can be used for tissue typing and in treating
 CC cancer. Anti-PRO antibodies can be used in diagnostic assays. AAF44270 to
 CC AAF44470 represent PCR primers and hybridisation probes used in the
 CC isolation of human PRO sequences. AAF44087 to AAF44269 and AAB65154 to
 CC AAB65300 represent human PRO polynucleotide and protein sequences given
 CC in the exemplification of the present invention

XX Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 4; Length 198;

Best Local Similarity 100.0%; Pred. No. 1.4e-104;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAKNFEDVRCKICPPYKENSCHIYKN 60
 DB 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAKNFEDVRCKICPPYKENSCHIYKN 60
 QY 61 ISQKDCDCLHVPEPMPVPGDVEAYCLRCCKYEERSSVTIKVTIIIVLSILGLLLYV 120
 DB 61 ISQKDCDCLHVPEPMPVPGDVEAYCLRCCKYEERSSVTIKVTIIIVLSILGLLLYV 120
 QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 QY 181 LQVQEQKSVFDRHVLS 198
 DB 181 LQVQEQKSVFDRHVLS 198

RESULT 11
 AAB61428

ID ABP61428 standard; protein; 198 AA.
XX AC ABP61428;
XX DT 30-SEP-2002 (first entry)
XX DE Human NF-kB activating protein SEQ ID NO 9.
XX Human; NF-kB; nuclear factor kappa B; mouse; antiinflammatory;
KW immunomodulator; cytostatic; antiinfective; osteopathic; nootropic;
KW neuroprotective; anti-HIV; autoimmune disease; cancer; infection;
KW bone disease; AIDS; neurodegenerative disease; ischaemic disorder.
XX OS Homo sapiens.
XX PN W0200253737-A1.
XX PD 11-JUL-2002.
XX PF 25-DEC-2001; 2001WO-JP011389.
XX PR 28-DEC-2000; 2000JP-0040288.
XX PR 26-MAR-2001; 2001JP-0008912.
XX PR 24-AUG-2001; 2001JP-00254018.
XX PA (ASAH) ASAH KASEI KOGYO KK.
XX PI Matsuda A, Honda G, Miyamatsu S, Nagano Y;
XX WPI; 2002-583617/62.
XX DR N-PSDB; ABQ91916.
XX PT NF-approximatelykB activating gene and expressed protein, applicable in
PT diagnosis and screening inhibitors or promoters to control excessive
PT activation or inhibition for treating e.g. inflammations, autoimmune
PT diseases and cancer.
XX PS Claim 1; Page 120-122; 841pp; Japanese.
XX CC The invention relates to a purified protein (I), comprising one of 90
CC fully defined sequences (ABP61424-ABP61513) or a protein based on any of
CC the sequences but with some amino acids deleted, substituted or added and
CC with a NF-kB (nuclear factor kappa B) activating effect. The protein and
CC encoding gene (ABQ91912-ABQ92001) are useful in diagnosis and screening
CC inhibitors or promoters to control excessive activation or inhibition and
CC for treating e.g. inflammations, autoimmune diseases, cancers,
CC infections, bone diseases, AIDS, neurodegenerative diseases or ischaemic
CC disorders
SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 5; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGLRLGSLISLALSLLVLLAQLSDAAKNFEDVRCKICPPYKENSHTYKN 60
DB 1 MATLWGLRLGSLISLALSLLVLLAQLSDAAKNFEDVRCKICPPYKENSHTYKN 60
QY 61 ISQKDCDCLHVVEPMPVGPDPVEAYCLRCCKEYERSVTIKVTIIYLSILGLLLYMW 120
DB 61 ISQKDCDCLHVVEPMPVGPDPVEAYCLRCCKEYERSVTIKVTIIYLSILGLLLYMW 120
QY 121 YLTLPVLPKRLFGHAQLIQSDDDIGDHPFANAHVLAERSRANVKNVEYAQRWK 180
DB 121 YLTLPVLPKRLFGHAQLIQSDDDIGDHPFANAHVLAERSRANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 12

ABU58100
ID ABU58100 standard; protein; 198 AA.
XX AC ABU58100;
XX DT 14-APR-2003 (first entry)
XX DE Human PRO polypeptide #132.
XX Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW horse; cow; dog; cat; sheep; pig; goat; rabbit; ADAPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX OS Homo sapiens.
XX PN US2003027163-A1.
XX PD 06-FEB-2003.
XX PF 15-NOV-2001; 2001US-00997666.
XX PR 16-JUN-1997; 97US-0049787P.
XX PR 17-OCT-1997; 97US-0062250P.
XX PR 05-NOV-1997; 97MO-US020069.
XX PR 12-NOV-1997; 97US-0065186P.
XX PR 13-NOV-1997; 97US-0065311P.
XX PR 24-NOV-1997; 97US-0066770P.
XX PR 25-FEB-1998; 98US-0075945P.
XX PR 20-MAR-1998; 98US-0078910P.
XX PR 28-APR-1998; 98US-0083322P.
XX PR 07-MAY-1998; 98US-0084600P.
XX PR 28-MAY-1998; 98US-0087106P.
XX PR 02-JUN-1998; 98US-0087607P.
XX PR 02-JUN-1998; 98US-0087609P.
XX PR 02-JUN-1998; 98US-0087759P.
XX PR 03-JUN-1998; 98US-0087827P.
XX PR 04-JUN-1998; 98US-0088021P.
XX PR 04-JUN-1998; 98US-0088025P.
XX PR 04-JUN-1998; 98US-0088036P.
XX PR 04-JUN-1998; 98US-0088038P.
XX PR 04-JUN-1998; 98US-0088039P.
XX PR 04-JUN-1998; 98US-0088030P.
XX PR 04-JUN-1998; 98US-0088033P.
XX PR 05-JUN-1998; 98US-0088167P.
XX PR 05-JUN-1998; 98US-0088202P.
XX PR 05-JUN-1998; 98US-0088212P.
XX PR 05-JUN-1998; 98US-0088217P.
XX PR 09-JUN-1998; 98US-0088655P.
XX PR 10-JUN-1998; 98US-0088734P.
XX PR 10-JUN-1998; 98US-0088738P.
XX PR 10-JUN-1998; 98US-0088742P.
XX PR 10-JUN-1998; 98US-0088810P.
XX PR 10-JUN-1998; 98US-0088824P.
XX PR 10-JUN-1998; 98US-0088826P.
XX PR 11-JUN-1998; 98US-0088858P.
XX PR 11-JUN-1998; 98US-0088861P.
XX PR 11-JUN-1998; 98US-0088876P.
XX PR 12-JUN-1998; 98US-0089105P.
XX PR 16-JUN-1998; 98US-0089440P.
XX PR 16-JUN-1998; 98US-0089512P.
XX PR 16-JUN-1998; 98US-0089514P.
XX PR 17-JUN-1998; 98US-0089532P.
XX PR 17-JUN-1998; 98US-0089538P.
XX PR 17-JUN-1998; 98US-0089598P.
XX PR 17-JUN-1998; 98US-0089599P.
XX PR 17-JUN-1998; 98US-0089600P.
XX PR 17-JUN-1998; 98US-0089653P.
XX PR 18-JUN-1998; 98US-0089601P.
XX PR 18-JUN-1998; 98US-0089907P.
XX PR 18-JUN-1998; 98US-0089908P.
XX PR 19-JUN-1998; 98US-0089947P.
XX PR 19-JUN-1998; 98US-0089948P.

```
PR 19-JUN-1998; 98US-0089552P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 23-JUN-1998; 98US-0090349P.
PR 23-JUN-1998; 98US-0090355P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090431P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090533P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090542P.
PR 24-JUN-1998; 98US-0090577P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 01-JUL-1998; 98US-0091360P.
PR 01-JUL-1998; 98US-0091344P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091519P.
PR 02-JUL-1998; 98US-0091526P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091633P.
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.
PR 07-JUL-1998; 98US-0091578P.
PR 07-JUL-1998; 98US-0091582P.
PR 09-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 20-JUL-1998; 98US-0093339P.
PR 30-JUL-1998; 98US-0094561P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095325P.
PR 10-AUG-1998; 98US-0095929P.
PR 10-AUG-1998; 98US-0096012P.
PR 11-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096146P.
PR 12-AUG-1998; 98US-0096329P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096950P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097218P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.

PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 21-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 16-SEP-1998; 98US-0100633P.
PR 17-SEP-1998; 98US-0100858P.
PR 17-SEP-1998; 98US-0100859P.
PR 07-OCT-1998; 98US-01021141.
PR 01-DEC-1998; 98US-01025108.
PR 22-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 98US-0113296P.
PR 08-MAR-1999; 98US-0113296P.
PR 12-MAR-1999; 98US-0123557P.
PR 12-MAR-1999; 98US-0123557P.
PR 23-JUN-1999; 98US-0141037P.
PR 07-JUL-1999; 98US-0143048P.
PR 26-JUL-1999; 98US-0144758P.
PR 26-JUL-1999; 98US-0145698P.
PR 28-JUL-1999; 98US-0146222P.
PR 17-AUG-1999; 98US-0149396P.
PR 15-SEP-1999; 98US-0149396P.
PR 15-SEP-1999; 98US-02021547.
PR 08-OCT-1999; 98US-0158663P.
PR 30-NOV-1999; 98US-0158663P.
PR 01-DEC-1999; 98US-0158663P.
PR 01-DEC-1999; 98US-0158663P.
PR 16-DEC-1999; 98US-0158663P.
PR 20-DEC-1999; 98US-0158663P.
PR 05-JAN-2000; 98US-0158663P.
PR 06-JAN-2000; 98US-0158663P.
PR 11-FEB-2000; 98US-0158663P.
PR 18-FEB-2000; 98US-0158663P.
PR 22-FEB-2000; 98US-0158663P.
PR 24-FEB-2000; 98US-0158663P.
PR 24-FEB-2000; 98US-0158663P.
PR 02-MAR-2000; 98US-0158663P.
PR 10-MAR-2000; 98US-0158663P.
PR 13-MAR-2000; 98US-0158663P.
PR 20-MAR-2000; 98US-0158663P.
PR 30-MAR-2000; 98US-0158663P.
PR 15-MAY-2000; 98US-0158663P.
PR 17-MAY-2000; 98US-0158663P.
PR 22-MAY-2000; 98US-0158663P.
PR 30-MAY-2000; 98US-0158663P.
PR 02-JUN-2000; 98US-0158663P.
PR 23-JUN-2000; 98US-0158663P.
PR 28-JUL-2000; 98US-0158663P.
PR 11-AUG-2000; 98US-0158663P.
PR 23-AUG-2000; 98US-0158663P.
PR 24-AUG-2000; 98US-0158663P.
PR 07-SEP-2000; 98US-0158663P.

Query Match 100.0%; Score 1031; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATLWGLRLGSLLSCLALSULLAQLSDAAKXNFEDVCKICPPYKENGSHIYKN 60
Db 1 MATLWGLRLGSLLSCLALSULLAQLSDAAKXNFEDVCKICPPYKENGSHIYKN 60
Qy 61 ISQKDCDCLHVVEPVPVGPVVEAYCLCECKYBERSSTVTKVIIIVLSILGLLLYV 120
Db 61 ISQKDCDCLHVVEPVPVGPVVEAYCLCECKYBERSSTVTKVIIIVLSILGLLLYV 120
Qy 121 YLTLVEPIKRLFGHAQLIOSDDDDIGHQPFANAHVLAERSRANVKNKVEAQRWK 180
Db 121 YLTLVEPIKRLFGHAQLIOSDDDDIGHQPFANAHVLAERSRANVKNKVEAQRWK 180
Qy 181 LQVQEQKSVFDRHVLS 198
```

Db 181 LQVQEQKSVFDRHVL 198
|||||
RESULT 13
ABU59178
ID ABU59178 standard; protein; 198 AA.
XX AC
XX AC ABU59178;
DT 28-APR-2003 (first entry)
XX DE
XX DE Novel human secreted or transmembrane protein PRO1375.
XX KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disorder;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis.
XX OS
XX OS Homo sapiens.
XX PN US2002132252-A1.
XX PD
XX PD 19-SEP-2002.
XX PF
XX PF 14-NOV-2001; 2001US-00990442.
PR 16-JUN-1997; 97US-0049787P.
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97WO-US020069.
PR 12-NOV-1997; 97US-0065186P.
PR 13-NOV-1997; 97US-0065311P.
PR 24-NOV-1997; 97US-0066770P.
PR 25-FEB-1998; 98US-0075945P.
PR 20-MAR-1998; 98US-0078910P.
PR 28-APR-1998; 98US-0083322P.
PR 07-MAY-1998; 98US-0084600P.
PR 28-MAY-1998; 98US-0087106P.
PR 02-JUN-1998; 98US-0087607P.
PR 02-JUN-1998; 98US-0087609P.
PR 03-JUN-1998; 98US-0087759P.
PR 04-JUN-1998; 98US-0088021P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088026P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088030P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088858P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.

PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089801P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 02-JUN-1999; 99WO-US012252.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 06-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 28-AUG-2001; 2001US-00941992.

(GETH) GENENTECH INC.

Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ,
Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF,
Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams FN, Wood WI,
Zhang Z;

WPI; 2003-247083/24.
N-PSDB; ABX80390.

Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
are therapeutically useful for enhancing immune response and in cancer
treatments.

Claim 12; Fig 300; 648pp; English.

The invention describes an isolated human PRO polypeptide. The PRO

RESULT 14	
ABU82690	
ID	ABU82690 standard; protein; 198 AA.
XX	
AC	ABU82690;
XX	
DT	26-JUN-2003 (first entry)
DE	Human secreted/transmembrane protein PRO1375.
XX	
XX	Human; PRO; secreted protein; transmembrane protein;
KW	cardiac insufficiency disorders; angiogenesis; wound healing;
KW	cancerous tumour; immune response; retinal disorder; sight loss;
KW	retinitis pigmentosum; age-related macular degeneration; AMD;
KW	kidney disorder; Berger disease; nephropathy; dermatitis; herpetiformis;
KW	Crohn's disease; sports injury; arthritis.
XX	
OS	Homo sapiens.
XX	
PN	US2003032023-A1.
XX	

PR	25-JUN-1998;	98US-0090578P.	PR	02-JUN-1999;	99WO-US012252.
PR	25-JUN-1998;	98US-0090690P.	PR	23-JUN-1999;	99US-0141037P.
PR	25-JUN-1998;	98US-0090694P.	PR	07-JUL-1999;	99US-0143048P.
PR	25-JUN-1998;	98US-0090695P.	PR	20-JUL-1999;	99US-0144758P.
PR	25-JUN-1998;	98US-0090866P.	PR	26-JUL-1999;	99US-0145698P.
PR	25-JUN-1998;	98US-0090862P.	PR	28-JUL-1999;	99US-0146222P.
PR	26-JUN-1998;	98US-0090863P.	PR	17-AUG-1999;	99US-0149396P.
PR	01-JUL-1998;	98US-0091360P.	PR	15-SEP-1999;	99WO-US021090.
PR	01-JUL-1998;	98US-0091544P.	PR	08-OCT-1999;	99WO-US021547.
PR	02-JUL-1998;	98US-0091478P.	PR	30-NOV-1999;	99WO-US028313.
PR	02-JUL-1998;	98US-0091519P.	PR	01-DEC-1999;	99WO-US028301.
PR	02-JUL-1998;	98US-0091626P.	PR	16-DEC-1999;	99WO-US028634.
PR	02-JUL-1998;	98US-0091628P.	PR	20-DEC-1999;	99WO-US030095.
PR	02-JUL-1998;	98US-0091633P.	PR	05-JAN-2000;	2000WO-US0030911.
PR	02-JUL-1998;	98US-0091646P.	PR	06-JAN-2000;	2000WO-US000219.
PR	02-JUL-1998;	98US-0091673P.	PR	11-FEB-2000;	2000WO-US000376.
PR	07-JUL-1998;	98US-0091978P.	PR	18-FEB-2000;	2000WO-US003565.
PR	09-JUL-1998;	98US-0092182P.	PR	22-FEB-2000;	2000WO-US004341.
PR	10-JUL-1998;	98US-0092472P.	PR	24-FEB-2000;	2000WO-US004914.
PR	20-JUL-1998;	98US-0093339P.	PR	02-MAR-2000;	2000WO-US005004.
PR	30-JUL-1998;	98US-0094651P.	PR	10-MAR-2000;	2000WO-US006319.
PR	04-AUG-1998;	98US-0095282P.	PR	15-MAR-2000;	2000WO-US006884.
PR	04-AUG-1998;	98US-0095285P.	PR	20-MAR-2000;	2000WO-US007377.
PR	04-AUG-1998;	98US-0095302P.	PR	30-MAR-2000;	2000WO-US008439.
PR	04-AUG-1998;	98US-0095318P.	PR	15-MAY-2000;	2000WO-US013358.
PR	04-AUG-1998;	98US-0095321P.	PR	17-MAY-2000;	2000WO-US013705.
PR	04-AUG-1998;	98US-0095325P.	PR	22-MAY-2000;	2000WO-US014042.
PR	10-AUG-1998;	98US-0095916P.	PR	30-MAY-2000;	2000WO-US014941.
PR	10-AUG-1998;	98US-0095929P.	PR	02-JUN-2000;	2000WO-US015264.
PR	10-AUG-1998;	98US-0096012P.	PR	23-JUN-2000;	2000US-0213637P.
PR	11-AUG-1998;	98US-0096143P.	PR	28-JUL-2000;	2000WO-US020710.
PR	11-AUG-1998;	98US-0096146P.	PR	11-AUG-2000;	2000WO-US022031.
PR	12-AUG-1998;	98US-0096329P.			
PR	17-AUG-1998;	98US-0096757P.			
PR	17-AUG-1998;	98US-0096766P.			
PR	17-AUG-1998;	98US-0096768P.			
PR	17-AUG-1998;	98US-0096773P.			
PR	17-AUG-1998;	98US-0096791P.			
PR	17-AUG-1998;	98US-0096867P.			
PR	17-AUG-1998;	98US-0096891P.			
PR	17-AUG-1998;	98US-0096884P.			
PR	17-AUG-1998;	98US-0096895P.			
PR	17-AUG-1998;	98US-0096897P.			
PR	18-AUG-1998;	98US-0096949P.			
PR	18-AUG-1998;	98US-0096950P.			
PR	18-AUG-1998;	98US-0096959P.			
PR	18-AUG-1998;	98US-0096960P.			
PR	18-AUG-1998;	98US-0097022P.			
PR	19-AUG-1998;	98US-0097141P.			
PR	20-AUG-1998;	98US-0097218P.			
PR	24-AUG-1998;	98US-0097661P.			
PR	26-AUG-1998;	98US-0097952P.			
PR	26-AUG-1998;	98US-0097954P.			
PR	26-AUG-1998;	98US-0097955P.			
PR	26-AUG-1998;	98US-0097971P.			
PR	26-AUG-1998;	98US-0097974P.			
PR	26-AUG-1998;	98US-0097978P.			
PR	26-AUG-1998;	98US-0097979P.			
PR	26-AUG-1998;	98US-0097986P.			
PR	26-AUG-1998;	98US-0098014P.			
PR	31-AUG-1998;	98US-0098525P.			
PR	16-SEP-1998;	98US-0100634P.			
PR	17-SEP-1998;	98WO-US019330.			
PR	17-SEP-1998;	98US-0100858P.			
PR	17-SEP-1998;	98WO-US019437.			
PR	07-OCT-1998;	98WO-US021141.			
PR	01-DEC-1998;	98WO-US025108.			
PR	22-DEC-1998;	98US-0113296P.			
PR	05-JAN-1999;	99WO-US000106.			
PR	08-MAR-1999;	99WO-US005028.			
PR	12-MAR-1999;	99US-0123357P.			
Query Match 100.0%; Score 1031; DB 6; Length 198;			Best Local Similarity 100.0%; Pred. No. 1.4e-104;		
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MATLWGGLRLGSLSLSLCLALSVLLLAQLSDAAKNFEDVRCKICIPYKENSCHIYNKN	60		
Db	1	MATLWGGLRLGSLSLSLCLALSVLLLAQLSDAAKNFEDVRCKICIPYKENSCHIYNKN	60		
QY	61	ISQKDCDCLHVPMPVGRGPDVEAYCLRCECKYEERSVTKVTIIIVLSILGLLLYV	120		
Db	61	ISQKDCDCLHVPMPVGRGPDVEAYCLRCECKYEERSVTKVTIIIVLSILGLLLYV	120		
QY	121	YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHDVLARSANVINKVEYAOQRWK	180		
Db	121	YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHDVLARSANVINKVEYAOQRWK	180		
QY	181	LQVQEQKSVFDRHVLS	198		
Db	181	LQVQEQKSVFDRHVLS	198		
RESULT 15					
ABU60609					
ID	ASU60609 standard; protein; 198 AA.				
XX	ABU60609;				
XX	01-MAY-2003 (first entry)				
DE	Human secreted/transmembrane protein, #168.				
XX	Human; PRO; secreted; transmembrane; signal peptide; pharmaceutical;				
KW	diagnostic; therapeutic; gene therapy.				
XX	Homo sapiens.				
XX	OS				
XX	US2002160384-A1.				
FN					

31-OCT-2002. 14-NOV-2001; 2001US-00992598. 17-JUN-1997; 97US-0049787P. 17-OCT-1997; 97US-0062250P. 05-NOV-1997; 97US-0020069. 12-NOV-1997; 97US-0065386P. 13-NOV-1997; 97US-0065311P. 24-NOV-1997; 97US-0066770P. 25-FEB-1998; 98US-0075945P. 20-MAR-1998; 98US-0078910P. 28-APR-1998; 98US-0083322P. 07-MAY-1998; 98US-0084600P. 28-MAY-1998; 98US-0087106P. 02-JUN-1998; 98US-0087607P. 02-JUN-1998; 98US-0087609P. 02-JUN-1998; 98US-0087759P. 03-JUN-1998; 98US-0087827P. 04-JUN-1998; 98US-0088021P. 04-JUN-1998; 98US-0088025P. 04-JUN-1998; 98US-0088026P. 04-JUN-1998; 98US-0088028P. 04-JUN-1998; 98US-0088029P. 04-JUN-1998; 98US-0088030P. 04-JUN-1998; 98US-0088033P. 04-JUN-1998; 98US-0088326P. 05-JUN-1998; 98US-0088167P. 05-JUN-1998; 98US-0088202P. 05-JUN-1998; 98US-0088212P. 05-JUN-1998; 98US-0088217P. 09-JUN-1998; 98US-0088655P. 10-JUN-1998; 98US-0088734P. 10-JUN-1998; 98US-0088738P. 10-JUN-1998; 98US-0088742P. 10-JUN-1998; 98US-0088810P. 10-JUN-1998; 98US-0088824P. 10-JUN-1998; 98US-0088836P. 11-JUN-1998; 98US-0088858P. 11-JUN-1998; 98US-0088861P. 11-JUN-1998; 98US-0088876P. 12-JUN-1998; 98US-0089105P. 16-JUN-1998; 98US-0089440P. 16-JUN-1998; 98US-0089512P. 16-JUN-1998; 98US-0089514P. 17-JUN-1998; 98US-0089532P. 17-JUN-1998; 98US-0089538P. 17-JUN-1998; 98US-0089598P. 17-JUN-1998; 98US-0089599P. 17-JUN-1998; 98US-0089600P. 17-JUN-1998; 98US-0089653P. 18-JUN-1998; 98US-0089801P. 18-JUN-1998; 98US-0089907P. 18-JUN-1998; 98US-0089908P. 16-SEP-1998; 98US-0019330. 17-SEP-1998; 98US-0019437. 07-OCT-1998; 98US-0021141. 01-DEC-1998; 98US-0025108. 05-JAN-1999; 98US-0000106. 08-MAR-1999; 98US-0005028. 02-JUN-1999; 98US-0012252. 15-SEP-1999; 98US-0021090. 15-SEP-1999; 98US-0021547. 30-NOV-1999; 98US-0028313. 01-DEC-1999; 98US-0028301. 01-DEC-1999; 98US-0028634. 16-DEC-1999; 98US-0030095. 20-DEC-1999; 98US-0030911. 05-JAN-2000; 2000US-0000219. 06-JAN-2000; 2000US-0000376. 11-FEB-2000; 2000US-0003565. 18-FEB-2000; 2000US-0004341.

31-OCT-2002. 14-NOV-2001; 2001US-00992598. 17-JUN-1997; 97US-0049787P. 17-OCT-1997; 97US-0062250P. 05-NOV-1997; 97US-0020069. 12-NOV-1997; 97US-0065386P. 13-NOV-1997; 97US-0065311P. 24-NOV-1997; 97US-0066770P. 25-FEB-1998; 98US-0075945P. 20-MAR-1998; 98US-0078910P. 28-APR-1998; 98US-0083322P. 07-MAY-1998; 98US-0084600P. 28-MAY-1998; 98US-0087106P. 02-JUN-1998; 98US-0087607P. 02-JUN-1998; 98US-0087609P. 02-JUN-1998; 98US-0087759P. 03-JUN-1998; 98US-0087827P. 04-JUN-1998; 98US-0088021P. 04-JUN-1998; 98US-0088025P. 04-JUN-1998; 98US-0088026P. 04-JUN-1998; 98US-0088028P. 04-JUN-1998; 98US-0088029P. 04-JUN-1998; 98US-0088030P. 04-JUN-1998; 98US-0088033P. 04-JUN-1998; 98US-0088326P. 05-JUN-1998; 98US-0088167P. 05-JUN-1998; 98US-0088202P. 05-JUN-1998; 98US-0088212P. 05-JUN-1998; 98US-0088217P. 09-JUN-1998; 98US-0088655P. 10-JUN-1998; 98US-0088734P. 10-JUN-1998; 98US-0088738P. 10-JUN-1998; 98US-0088742P. 10-JUN-1998; 98US-0088810P. 10-JUN-1998; 98US-0088824P. 10-JUN-1998; 98US-0088836P. 11-JUN-1998; 98US-0088858P. 11-JUN-1998; 98US-0088861P. 11-JUN-1998; 98US-0088876P. 12-JUN-1998; 98US-0089105P. 16-JUN-1998; 98US-0089440P. 16-JUN-1998; 98US-0089512P. 16-JUN-1998; 98US-0089514P. 17-JUN-1998; 98US-0089532P. 17-JUN-1998; 98US-0089538P. 17-JUN-1998; 98US-0089598P. 17-JUN-1998; 98US-0089599P. 17-JUN-1998; 98US-0089600P. 17-JUN-1998; 98US-0089653P. 18-JUN-1998; 98US-0089801P. 18-JUN-1998; 98US-0089907P. 18-JUN-1998; 98US-0089908P. 16-SEP-1998; 98US-0019330. 17-SEP-1998; 98US-0019437. 07-OCT-1998; 98US-0021141. 01-DEC-1998; 98US-0025108. 05-JAN-1999; 98US-0000106. 08-MAR-1999; 98US-0005028. 02-JUN-1999; 98US-0012252. 15-SEP-1999; 98US-0021090. 15-SEP-1999; 98US-0021547. 30-NOV-1999; 98US-0028313. 01-DEC-1999; 98US-0028301. 01-DEC-1999; 98US-0028634. 16-DEC-1999; 98US-0030095. 20-DEC-1999; 98US-0030911. 05-JAN-2000; 2000US-0000219. 06-JAN-2000; 2000US-0000376. 11-FEB-2000; 2000US-0003565. 18-FEB-2000; 2000US-0004341.

31-OCT-2002. 14-NOV-2001; 2001US-00992598. 17-JUN-1997; 97US-0049787P. 17-OCT-1997; 97US-0062250P. 05-NOV-1997; 97US-0020069. 12-NOV-1997; 97US-0065386P. 13-NOV-1997; 97US-0065311P. 24-NOV-1997; 97US-0066770P. 25-FEB-1998; 98US-0075945P. 20-MAR-1998; 98US-0078910P. 28-APR-1998; 98US-0083322P. 07-MAY-1998; 98US-0084600P. 28-MAY-1998; 98US-0087106P. 02-JUN-1998; 98US-0087607P. 02-JUN-1998; 98US-0087609P. 02-JUN-1998; 98US-0087759P. 03-JUN-1998; 98US-0087827P. 04-JUN-1998; 98US-0088021P. 04-JUN-1998; 98US-0088025P. 04-JUN-1998; 98US-0088026P. 04-JUN-1998; 98US-0088028P. 04-JUN-1998; 98US-0088029P. 04-JUN-1998; 98US-0088030P. 04-JUN-1998; 98US-0088033P

Db	61	ISQKDCDLHVBFMPVRGPDVEAYCLRCCKYEERSVTIKVTIIYLSILGLLLLYMV	120
Qy	121	YLTIVEPILKRLFGHAQLIQSDDDIGDHQPFANAHDLVLRSSRANVLNKVEYAOQRWK	180
Db	121	YLTIVEPILKRLFGHAQLIQSDDDIGDHQPFANAHDLVLRSSRANVLNKVEYAOQRWK	180
Qy	181	LQVQQRKSVFDRHVLS	198
Db	181	LQVQQRKSVFDRHVLS	198

Search completed: March 15, 2004, 08:25:47
Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:20:07 ; Search time 18 Seconds
(without alignments)
572.772 Million cell updates/sec

Title: US-09-997-573-418

Perfect score: 1031

Sequence: 1 MATLWGGLRLGSLLSCL.....WKLQVQKRSVDRHVLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	100.0	198	1	CK15_HUMAN
2	979.5	95.0	199	1	Q9nq34 homo sapien
3	544	52.8	183	1	CK15_MOUSE
4	541.5	52.5	183	1	Q9p0t7 homo sapien
5	530	51.4	187	1	Q9cr23 mus musculus
6	226.5	22.0	227	1	Q73698 fugu rubrip
7	81.5	7.9	606	1	Q9vna4 drosophila
8	80.5	7.8	895	1	Q89016 mus musculus
9	79	7.7	642	1	P55850 mus musculus
10	78.5	7.6	1700	1	Q8fp91 corynebacte
11	78	7.6	285	1	Q03376 chironomus
12	77.5	7.5	245	1	P75098 mycoplasma
13	77.5	7.5	318	1	Q63345 rattus norv
14	77.5	7.5	1257	1	Q9zzz1 hippopotamu
15	77.5	7.5	5703	1	P06494 rattus norv
16	76	7.4	318	1	Q9hc84 homo sapien
17	75.5	7.3	315	1	O78696 lemur catta
18	75.5	7.3	318	1	P03888 mus musculus
19	75.5	7.3	402	1	O78699 tamandua te
20	75	7.3	358	1	P39542 saccharomyc
21	74.5	7.2	318	1	P3280 neurospora
22	74.5	7.2	1630	1	Q9xy4 papio hamad
23	74	7.2	246	1	Q3018 saccharomyc
24	74	7.2	256	1	P5803 bos taurus
25	74	7.2	606	1	P20334 mus musculus
26	74	7.2	1002	1	O14678 homo sapien
27	74	7.2	1304	1	P29815 ansacta moo
28	73.5	7.1	1958	1	P08575 homo sapien
29	73	7.1	318	1	O60152 schizosacch
30	73	7.1	896	1	O78706 phascoglyc
31	72.5	7.0	303	1	Q28060 bos taurus
32	72.5	7.0	318	1	Y247_HUMAN
33	72.5	7.0	318	1	P41296 balaenopter
					P24967 balaenopter

34	72	7.0	584	1	RECN_SVNY3
35	72	7.0	1207	1	EGF_HUMAN
36	71.5	6.9	318	1	NU1M_BRAVA
37	71.5	6.9	323	1	NU1M_PARLI
38	71.5	6.9	1264	1	CYAS_RABIT
39	71.5	6.9	2482	1	VWF_PIG
40	70.5	6.8	1064	1	YY08_METJA
41	70.5	6.8	1194	1	MGRI_HUMAN
42	70	6.8	318	1	NU1M_MONDO
43	70	6.8	381	1	COS8_YEAST
44	70	6.8	495	1	PA18_HUMAN
45	70	6.8	720	1	YD71_SCHPO

ALIGNMENTS

RESULT 1

CK15_HUMAN					
ID	CK15_HUMAN	STANDARD;	PRT;	198 AA.	
AC	Q9NQ34;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Protein Cllorf15.				
GN	Cllorf15.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21418998; PubMed=11528127;				
RA	Amid C., Bahr A., Mujica A., Sampson N., Bikar S.E., Winterpacht A.,				
RA	Zabel B., Hankeln T., Schmidt E.R.;				
RT	"Comparative genomic sequencing reveals a strikingly similar				
RT	architecture of a conserved syntenic region on human chromosome				
RT	11p15.3 (including gene ST5) and mouse chromosome 7.";				
RT	Cytogenet. Cell Genet. 93:284-290(2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
TI	TISSUE=Pancreas;				
RC	MEDLINE=22388257; PubMed=12477932;				
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalón D.K., Musny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,				
RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.,				
RT	"Generation and initial analysis of more than 15,000 full-length				
RT	human and mouse cDNA sequences.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
CC	-1- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).				
CC	-1- SIMILARITY: Belongs to the TMEM9 family.				

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).


```
CC -----
DR EMBL; AJ400877; CAB92287.1; -
DR EMBL; BC040124; AAH40124.1; -
DR Genew; HGNC:1168; C11orf15.
DR InterPro; IPR004153; CXCC repeat.
DR InterPro; IPR008853; TMEM9.
DR Pfam; PF03128; CXCC; 1.
DR Pfam; PF05434; TMEM9; 1.
DR Transmembrane.
KW Transmembrane.
FT TRANSMEM 103 125
SQ SEQUENCE 198 AA; 22531 MW; 888E5536A01B44F CRC64;

Query Match 100.08; Score 1031; DB 1; Length 198;
Best Local Similarity 100.08; Pred. No. 4e-94;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGSLRLGSLLSCLALSLLVLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNNK 60
DB 1 MATLWGGSLRLGSLLSCLALSLLVLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNNK 60
QY 61 ISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEKSSVTIKVITIIYLSILGLLLYV 120
DB 61 ISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEKSSVTIKVITIIYLSILGLLLYV 120
QY 121 YLTLPVEPKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLNKVEYAQQRWK 180
DB 121 YLTLPVEPKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLNKVEYAQQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198

RESULT 2
CK15_MOUSE STANDARD; PRT; 199 AA.
AC Q9JUR8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein C11orf15 homolog.
GN C11ORF15
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21418998; PubMed=11528127;
RA Amid C., Bahr A., Mujica A., Sampson N., Bikar S.E., Winterpacht A.,
Zabel B., Hankeln T., Schmidt E.R.;
RT "Comparative genomic sequencing reveals a strikingly similar
architecture of a conserved syntenic region on human chromosome
11p15.3 (including gene ST5) and mouse chromosome 7.";
RL Cytogenet. Cell Genet. 93:284-290(2001).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Pancreas, and Tongue;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bonc H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schröter L.M., Scubili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
```

```
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=2338257; PubMed=12477932;
RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lomellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards J., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RN SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC - SIMILARITY: Belongs to the TMEM9 family.
[5]
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
[6]
CC EMBL; AJ400878; CAB92295.1; -
CC EMBL; AK009147; BAB26105.1; -
CC EMBL; AK007430; BAB25033.1; -
CC EMBL; BC022208; AAH02208.1; -
CC InterPro; IPR004153; CXCC repeat.
CC InterPro; IPR008853; TMEM9.
CC Pfam; PF03128; CXCC; 1.
CC Pfam; PF05434; TMEM9; 1.
CC Transmembrane.
KW TRANSMEM 12 34
FT TRANSMEM 104 126
SQ SEQUENCE 199 AA; 22607 MW; 476B6FD1BCECA4E5 CRC64;

Query Match 95.08; Score 979.5; DB 1; Length 199;
Best Local Similarity 94.54; Pred. No. 4.6e-89;
Matches 188; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 MATLW-GGLRLGSLLSCLALSLLVLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNNK 59
DB 1 MASLWCGNLLRLGSLLSCLALSLLVLAQLTGAKNFEDVRCCKICPPYKENSCHIYNNK 60
QY 60 NISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEKSSVTIKVITIIYLSILGLLLYV 119
DB 61 NISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEKSSVTIKVITIIYLSILGLLLYV 120
QY 120 VYLTLPVEPKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLNKVEYAQQRW 179
DB 121 VYLTLPVEPKRLFGHSQLQSDDDVGDHPFANAHDVLAARSRAVNLNKVEYAQQRW 180
QY 180 KLVQVQQRKSVFDRHVLS 198
DB 181 KLVQVQQRKSVFDRHVLS 199
```

RESULT 3

TM9 HUMAN

ID NAME STANDARD; PRT; 183 AA.

AC Q9P0T7, Q96N09; Q96QF5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Transmembrane protein 9 precursor (dermal papilla derived protein 4)

DE (Protein HSPC186).

GN TMEM9 OR DERP4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI TaxID=9606;

OX [1]

SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.

RP MEDLINE=22946083; PubMed=12359240;

RP Kveine M., Tenstad E., Dosen G., Funderud S., Rian E.;

RA "Characterization of the novel human transmembrane protein 9 (TMEM9)

RT that localizes to lysosomes and late endosomes.";

RL Biochem. Biophys. Res. Commun. 297:912-917(2002).

RL [2]

SEQUENCE FROM N.A.

RP TISSUE=dermal papilla;

RP Ikeda A., Yoshimoto M.;

RA "Molecular cloning of a dermal papilla derived gene.";

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RL [3]

SEQUENCE FROM N.A.

RP TISSUE=Blood;

RP MEDLINE=20499367; PubMed=11042152;

RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.;

RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,

RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;

RA "Cloning and functional analysis of cDNAs with open reading frames for

RT 300 previously undefined genes expressed in CD34+ hematopoietic

RT stem/progenitor cells.";

RL Genome Res. 10:1546-1560(2000).

RL [4]

SEQUENCE FROM N.A.

RP TISSUE=Brain;

RP MEDLINE=21154917; PubMed=11230166;

RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,

RA Ansoerge W., Boecker M., Blecker H., Baurerachs S., Blum H.,

RA Lauber J., Dueterhoeft A., Beyer A., Koehler K., Strack N.,

RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,

RA Wambutt R., Korn B., Klein M., Poustka A.;

RA "Towards a catalog of human genes and proteins: sequencing and

RT analysis of 500 novel complete protein coding human cDNAs.";

RL Genome Res. 11:422-435(2001).

RL [5]

SEQUENCE FROM N.A.

RP TISSUE=Cerebellum;

RP Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,

RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,

RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,

RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,

RA Sugano S., Nagahori K., Masuho Y., Nagai K., Isogai T.;

RA "NEDO human cDNA sequencing project.";

RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RL [6]

SEQUENCE FROM N.A.

RP Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,

RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,

RA Nagahori K., Sugano S., Isogai T.;

RT "HRI human cDNA sequencing project.";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RL [7]

SEQUENCE FROM N.A.

RP TISSUE=Kidney;

RA [8]

RA [9]

RA [10]

RA [11]

RA [12]

RA [13]

RA [14]

RA [15]

RA [16]

RA [17]

RA [18]

RA [19]

RA [20]

RA [21]

RA [22]

RA [23]

RA [24]

RA [25]

RA [26]

RA [27]

RA [28]

RA [29]

RA [30]

RA [31]

RA [32]

RA [33]

RA [34]

RA [35]

RA [36]

RA [37]

RA [38]

RA [39]

RA [40]

RA [41]

RA [42]

RA [43]

RA [44]

RA [45]

RA [46]

RA [47]

RA [48]

RA [49]

RA [50]

RA [51]

RA [52]

RA [53]

RA [54]

RA [55]

RA [56]

RA [57]

RA [58]

RA [59]

RA [60]

RA [61]

RA [62]

RA [63]

RA [64]

RA [65]

RA [66]

RA [67]

RA [68]

RA [69]

RA [70]

RA [71]

RA [72]

RA [73]

RA [74]

RA [75]

RA [76]

RA [77]

RA [78]

RA [79]

RA [80]

RA [81]

RA [82]

RA [83]

RA [84]

RA [85]

RA [86]

RA [87]

RA [88]

RA [89]

RA [90]

RA [91]

RA [92]

RA [93]

RA [94]

RA [95]

RA [96]

RA [97]

RA [98]

RA [99]

RA [100]

RA [101]

RA [102]

RA [103]

RA [104]

RA [105]

RA [106]

RA [107]

RA [108]

RA [109]

RA [110]

RA [111]

RA [112]

RA [113]

RA [114]

RA [115]

RA [116]

RA [117]

RA [118]

RA [119]

RA [120]

RA [121]

RA [122]

RA [123]

RA [124]

RA [125]

RA [126]

RA [127]

RA [128]

RA [129]

RA [130]

RA [131]

RA [132]

RA [133]

RA [134]

RA [135]

RA [136]

RA [137]

RA [138]

RA [139]

RA [140]

RA [141]

RA [142]

RA [143]

RA [144]

RA [145]

RA [146]

RA [147]

RA [148]

RA [149]

RA [150]

RA [151]

RA [152]

RA [153]

RA [154]

RA [155]

RA [156]

RA [157]

RA [158]

RA [159]

RA [160]

RA [161]

RA [162]

RA [163]

RA [164]

RA [165]

RA [166]

RA [167]

RA [168]

RA [169]

RA [170]

RA [171]

RA [172]

RA [173]

RA [174]

RA [175]

RA [176]

RA [177]

RA [178]

RA [179]

RA [180]

RA [181]

RA [182]

RA [183]

RA [184]

RA [185]

RA [186]

RA [187]

RA [188]

RA [189]

RA [190]

RA [191]

RA [192]

RA [193]

RA [194]

RA [195]

RA [196]

RA [197]

RA [198]

RA [199]

RA [200]

RA [201]

RA [202]

RA [203]

RA [204]

RA [205]

RA [206]

RA [207]

RA [208]

RA [209]

RA [210]

RA [211]

RA [212]

RA [213]

RA [214]

RA [215]

RA [216]

RA [217]

RA [218]

RA [219]

RA [220]

RA [221]

RA [222]

RA [223]

RA [224]

RA [225]

RA [226]

RA [227]

RA [228]

RA [229]

RA [230]

RA [231]

RA [232]

RA [233]

RA [234]

RA [235]

RA [236]

RA [237]

RA [238]

RA [239]

RA [240]

RA [241]

RA [242]

RA [243]

RA [244]

RA [245]

RA [246]

RA [247]

RA [248]

RA [249]

RA [250]

RA [251]

RA [252]

RA [253]

RA [254]

RA [255]

RA [256]

RA [257]

RA [258]

RA [259]

RA [260]

RA [261]

RA [262]

RA [263]

RA [264]

RA [265]

RA [266]

RA [267]

RA [268]

RA [269]

RA [270]

RA [271]

RA [272]

RA [273]

RA [274]

RA [275]

RA [276]

RA [277]

RA [278]

RA [279]

RA [280]

RA

RA Riboldi,Tunnicliffe G.R., Platzer M., Nyakatura G., Elgar G.S.,
RA Brenner S., Rosenthal A.,
RA "Analysis of the genomic loci of Fugu rubripes homologs of the human
RT disease genes L1CAM, G6PD and F55";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC -!- SIMILARITY: Belongs to the TMEM9 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF026198; AAC15584.1; -
DR PRF: T30536; T30536.
DR GO: GO:0005770; C:late endosome; ISS.
DR GO: GO:0005764; C:lysosome; ISS.
DR InterPro: IPR004153; CXCXC_repeat.
DR InterPro: IPR008853; TMEM9.
DR Pfam: PF03128; CXCXC; 1.
DR Pfam: PF05434; TMEM9; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 10 27 POTENTIAL.
FT TRANSMEM 99 121 POTENTIAL.
FT NON TPR 187 187
SQ SEQUENCE 187 AA; 21460 MW; 51DD101FE4666D3B CRC64;
Query Match 51.48; Score 530; DB 1; Length 187;
Best Local Similarity 57.88; Pred. No. 6.3e-45;
Matches 108; Conservative 29; Mismatches 40; Indels 10; Gaps 4;
QY 14 LLSLSIALSVLLLAQLSDA--AKNFDVRCKICPPYKENSCHIYKNIKSKDCDLHV 71
DB 8 LMLAALPMTLFLDLNVGVVQAQSFDDVRCKICPPYRNISCHIYKNIKSKDCNCLHV 67
QY 72 VEPNVEGPDVEAYCLCECKYERSVTKVITIIYVLTGLLLYMYVLTVEPIKR 131
DB 68 VDPNVEGPDVEAYCLCECKYERSVTKVITIIYVLTGLLLYMYVLTVEPIKR 126
QY 132 RLFGHAQLIQSDDIGHQPFANAHDLARSRRAN-VLNKVEYAQQRKLVQEQKSV 190
DB 127 KPDLAQLTHNEEDSEDIQ-----QMSGDPARGNTVLERVEGAQQRKLVQEQKSV 180
QY 191 FDRHVL 197
DB 181 FDRHVL 187
RESULT 6
ID YB61_DROME STANDARD; PRT; 227 AA.
AC Q9VNA4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein CG1161.
GN CG1161.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Fertaz C., Ferrera S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
CC -!- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC -!- SIMILARITY: Belongs to the TMEM9 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003603; AAF52042.1; -
DR FlyBase: FBgn0037313; CG1161.
DR InterPro: IPR008853; TMEM9.
DR Pfam: PF05434; TMEM9; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 7 24 POTENTIAL.
FT TRANSMEM 135 157 POTENTIAL.
SQ SEQUENCE 227 AA; 24653 MW; 7B3F0B42B0A851EE CRC64;
Query Match 22.04; Score 226.5; DB 1; Length 227;
Best Local Similarity 31.64; Pred. No. 5.2e-15;
Matches 50; Conservative 38; Mismatches 63; Indels 7; Gaps 4;
QY 43 KCICPPYKENSCHIYKNIKSKDCDLHVPEP-MVVRGPDVEAYCLCECKYERSVTK 101
DB 75 KVIGQPVLAAPGKNSNSSTTECVACAGALLPRLDANGKEL-PICAECKSHVARTTLI 133
QY 102 KVTIIYVLTGLLLYMYVLTVEPIKR-LFGHAQLIQSDDIGHQPFANAHDLA 160
DB 134 KVVIIIVITISILYVLTFLMCLDFLLNKRKVKANYQETHNEDDEPTPLPAVNNQELSA 193
QY 161 RSRANVLNKVEYAQQRKLVQEQKSVFDRHVLVS 198
DB 194 ----RANVLNRVGHQDKWKQVRQRRHYDRHTMLN 227
RESULT 7
ABD4_MOUSE

ID ABD4_MOUSE STANDARD; PRT; 606 AA.
AC O89016;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATP-binding cassette, sub-family D, member 4 (Peroxisomal membrane
protein 69) (PMP69) (Peroxisomal membrane protein 1-like) (PMP1-L)
(P70R).
GN ABCD4 OR PMP1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98409337; PubMed=9738957;
RA Holzinger A., Muntau A., Mayerhofer P., Kammerer S., Albet S.,
RT "The mouse gene encoding the peroxisomal membrane protein 1-like
protein (PMP1-L): cDNA cloning, genomic organization and comparative
expression studies."
RL FEBS Lett. 433:179-183 (1998).
CC -1- SUBUNIT: Homodimer or heterodimer (Potential).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -1- SIMILARITY: Belongs to the ABC transporter family. ABC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ001166; CAA04570.1; --
DR MGD; MGI:1349217; Abcd4.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00249; ABC_TM1F; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Peroxisome.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
FT NP BIND 421 428 ATP (POTENTIAL).
SQ SEQUENCE 606 AA; 68549 MW; 1EF6F044D665A200 CRC64;
Query Match 7.9%; Score 81.5; DB 1; Length 606;
Best Local Similarity 23.8%; Pred. No. 2.7;
Matches 41; Conservative 25; Mismatches 57; Indels 49; Gaps 8;
QY 7 GLRLGLSLGLSLALSVLLLAQSDAANKFEDVRCKICPPYK-----ENSGHIYNK--- 59
DB 69 GVLGKDLGKALTLLAVTLVLNGLSTKSFQFTCNLLVSVWRKDLTEHLHLYFRARV 128
QY 60 -----NLSQKDCDCLHVPEMFVPGDVAFLRCECKVEERSVTKV-----TIIIV-- 108
DB 129 YITLVLRDID-----NPQRTISQDVRFPC-----RLQSSVTSKLIISFTIYYTY 176
QY 109 -----LSILGLLLL-YMVYLTLPVLPRLFGLHAQLIQSDDDIGD 148
DB 177 QCFQSTGWLGPVIFGYFVGTVMVNTKLTMPIV-----TKLVQCKLEGD 221
RESULT 8
DSC3_MOUSE
ID DSC3_MOUSE STANDARD; PRT; 895 AA.

P55950; O55110; O55122;
01-NOV-1997 (Rel. 35, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Desmocollin 3 precursor.
GN DSC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=98049324; PubMed=9389456;
RA Chidsey M.A.J., Yue K.K.M., Gould S., Byrne C., Garrod D.R.;
RT "Changing pattern of desmocollin 3 expression accompanies epidermal
organisation during skin development."
RL Dev. Dyn. 210:315-327 (1997).
RN [2]
RP SEQUENCE OF 709-874 FROM N.A., ALTERNATIVE SPLICING, AND TISSUE
SPECIFICITY.
RX MEDLINE=98067789; PubMed=9404003;
RA King I.A., Angst B.D., Hunt D.M., Kruger M., Arnmann J., Buxton R.S.;
RT "Hierarchical expression of desmosomal cadherins during stratified
epithelial morphogenesis in the mouse."
RL Differentiation 62:83-96 (1997).
CC -1- FUNCTION: Component of intercellular desmosome junctions. Involved
in the interaction of plaque proteins and intermediate filaments
mediating cell-cell adhesion. May contribute to epidermal cell
positioning (stratification) by mediating differential
adhesiveness between cells that express different isoforms.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=3A;
ISOId=P55950-1; Sequence=Displayed;
Names=3B;
ISOId=P55950-2; Sequence=VSP_000665; VSP_000666;
TISSUE SPECIFICITY: First expressed at E13.0 in epithelium of
whisker pads and external nares, and in most mature vibrissa
follicles. 12 hours later, prominently expressed in whiskers and
tactile follicles above the eye. At E14.5, also expressed in
developing nails and teeth and, at low levels, in ventral and
lateral skin. At E15.5, highly expressed in general body epidermis
and at E16.5, detected over entire embryo. In the adult, highly
expressed in basal layers of stratified cells.
-1- DOMAIN: Calcium may be bound by the cadherin-like repeats
(Potential).
-1- SIMILARITY: Contains 5 cadherin domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; Y11169; CAA72045.1; --
EMBL; AJ000329; CAA03996.1; --
HSSP; P15116; INCJ.
MGD; MGI:1194993; Dsc3.
InterPro; IPR002126; Cadherin.
InterPro; IPR00233; Cadherin_C_term.
Pfam; PF00028; cadherin; 5.
Pfam; PF01049; Cadherin_C_term; 1.
PRINTS; PR00205; CADHERIN.
SMART; SM00112; CA; 5.
PROSITE; PS00232; CADHERIN_1; 3.
PROSITE; PS00268; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
Cytoskeleton; Calcium-binding; Alternative splicing.
FT SIGNAL 1 31 POTENTIAL.
FT PROPEP 32 135 POTENTIAL.

```

FT CHAIN 136 895 DESMOCOLLIN 3.
FT DOMAIN 136 695 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 696 716 POTENTIAL.
FT DOMAIN 717 895 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 136 243 CADHERIN 1.
FT DOMAIN 244 355 CADHERIN 2.
FT DOMAIN 356 472 CADHERIN 3.
FT DOMAIN 473 580 CADHERIN 4.
FT DOMAIN 581 691 CADHERIN 5.
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 831 838 KLHVCNQN -> DSIRHTG (in isoform 3B).
FT VARSPLIC 839 895 /FtId=VSP_000665.
FT VARSPLIC 839 895 Missing (in isoform 3B).
FT VARSPLIC 839 895 /FtId=VSP_000666.
FT CONFLICT 715 715 V -> I (IN REF. 2).
FT CONFLICT 750 767 WHPHPCQQLPQLLMTN -> NGPHTTANNSSQGFCTGM
FT CONFLICT 750 767 (IN REF. 2).
FT CONFLICT 840 840 N -> D (IN REF. 2).
FT CONFLICT 873 874 EG -> IK (IN REF. 2).
SQ SEQUENCE 895 AA; 100653 MW; D561E67C57232270 CRC64;

Query Match 7.8%; Score 80.5; DB 1; Length 895;
Best Local Similarity 25.7%; Pred. No. 5.2; Indels 53; Gaps 5;
Matches 37; Conservative 17; Mismatches 53;

QY 29 QLSAAKNFE-----DVRCKICPPYKENGSHYNNKISQKDCDCLHVVPEMPVRGPDVE 83
DB 628 QVNTAARLSVQKTADVIQVNPVTVDRAGQSATKILRVNLCDIH----- 674
QY 84 AYCLRCCKYBEERSVTIKVTIIYLSI-LGLLLLYMVLTLVEPIKRLRFOHAQLIOS 142
DB 675 ----PSQCPLRSRA-GITLGKWAYLAILGIALLFSLVLLTVGVVTAKE----- 720
QY 143 DDDIGDHOPFANADVLARSRA 166
DB 721 ----GKFFPEDLAHENLIISNTA 740

RESULT 9
MQO COREF
ID MQO COREF STANDARD; PRT; 642 AA.
AC Q8PF91;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable malate:quinone oxidoreductase (EC 1.1.99.16) (Malate
DE dehydrogenase [acceptor]) (MQO).
GN MQO OR CE1894.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=27273752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
CC -!- CATALYTIC ACTIVITY: (S)-malate + acceptor = oxaloacetate + reduced
CC acceptor.
CC -!- COFACTOR: FAD (By similarity).
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SIMILARITY: In the C-terminal section; belongs to the MQO family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to a license@isb-sib.ch).
CC -----
CC EMBL; AP005220; BAC18704.1; -
CC HAMAP; MF 00212; fused; 1.
CC InterPro; IPR006231; mal_quin_oxido; 1.
CC TIGRPFAMs; TIGR01320; mal_quin_oxido; 1.
CC Oxidoreductase; Tricarboxylic acid cycle; Flavoprotein; FAD;
KW Complete proteome.
KW DOMAIN 1 142 UNKNOWN.
FT DOMAIN 143 642 MQO.
SQ SEQUENCE 642 AA; 70125 MW; 19E8271DBF472725 CRC64;

Query Match 7.7%; Score 79; DB 1; Length 642;
Best Local Similarity 22.5%; Pred. No. 5.1; Indels 64; Gaps 11;
Matches 42; Conservative 26; Mismatches 55;

QY 39 DVRCKICPPYKENGSHYNNKISQKDCDCLHVVPEMPVRGPDVE-----AYCLRCCKY 93
DB 87 ELRTTCTCTPSRSRPGQVFT-----TCRSAIH--PCAGPRVDSCVQHAGCVRTTTRG 136
QY 94 EER-----SSVTIKVTII-----IYLSILGLLLYMYVTLVEP-----ILKRL 133
DB 137 DKRKLMSDSPKQKQVTDADVIVGVGINSAILGMLRL-----EPWSQVIFERL 190
QY 134 FOHAQLIOSDDIDGHPFPANA---HDVLAISR-----SRA-NVLNKVEYAQR 178
DB 191 DGAQCESS-----PMNAGTGHSAICELNVTPEKNGKIDISKAVNINEKTVSRQF 242
QY 179 WKLQVQE 185
DB 243 WAHVEE 249

RESULT 10
BAR3 CHITE
ID BAR3 CHITE STANDARD; PRT; 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Balbiani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Widge).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;
OC Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1699777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RT repetitive structure split by many introns."
RL J. Mol. Biol. 211:331-349(1990).
CC -!- FUNCTION: Used by the larvae to construct a supramolecular
CC structure, the larval tube. Balbiani ring protein 3 could play a
CC role as a transport protein that binds to other proteins
CC intracellularly and in the gland lumen in order to prevent these
CC from forming water-insoluble fibers too early.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Salivary gland.
CC -!- DOMAIN: Has 82 approximate repeats of Cys-x-Cys-x-Cys.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X52263; CRA36506.1; --
DR PIR; S08157; S08167.
DR HSP; P15358; 18KZ.
DR InterPro; IPR004153; CXKC_repeat.
DR Pfam; PF03128; CXKC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;
Query Match 7.6%; Score 78.5; DB 1; Length 1700;
Best Local Similarity 24.5%; Pred. No. 17;
Matches 23; Conservative 9; Mismatches 39; Indels 23; Gaps 2;
QY 35 KNFEDVRCKICPPYK-----ENSGHYNNKXISQKDCDLHV 72
DB 898 KFNDFTCSCGSPGKLDCTGNTKWSAETCTCGGDVNRNGCNKFNDFNLCCQCEKNKQ 957
QY 73 EMPVRGPDVEAY-CLRCECKVEERSSVTIKVTI 105
DB 958 EVANCKSPRTWNYDTCCKVCNADSDDCVRFQI 991

RESULT 11

ID Y011_MYCPN STANDARD; PRT; 285 AA.
AC P75098;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MG011 homolog (D12_orf285).
GN MPN015 OR MP139.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / ML29;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelsreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449 (1996).
CC -1- SIMILARITY: BELONGS TO THE ATP-GRASP SUPERFAMILY. RIMK FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE000016; AAB95787.1; --
DR PIR; S73465; S73465.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 285 AA; 33446 MW; 1D0F653B3EA5049F CRC64;
Query Match 7.6%; Score 78; DB 1; Length 285;
Best Local Similarity 28.6%; Pred. No. 2.6;
Matches 22; Conservative 12; Mismatches 23; Indels 20; Gaps 3;

QY 124 LVEPILKRFLFGHQLIQSDDDTDGHPFANADHVLRSP-----SRANVLKV----- 172
DB 30 LVDELNKKVLGHVILL--DDETADHKHKIKVELIINRSRIDFLTKHFLNFSLINPON 87
QY 173 -----EYAQQRWKLQ 182

Db 88 IVLVANDKYETRWLQK 104

RESULT 12

MOG_RAT
ID MOG_RAT STANDARD; PRT; 245 AA.
AC Q63345;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93085763; PubMed=1453482;
RA Gardiner M.V., Amiguet P., Linington C., Matthieu J.-M.;
RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
RT immunoglobulin superfamily";
RL J. Neurosci. Res. 33:177-187 (1992).
RN [2]
RP SEQUENCE OF 28-245 FROM N.A.
RC TISSUE=Brain; PubMed=8367453;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Roussel G.,
RA Pontarotti P., Roessel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994 (1993).
RN [3]
RP STRUCTURE BY NMR OF G2-82.
RX MEDLINE=97354172; PubMed=9210466;
RA Albouze-Abou S., Wilson J.C., Bernard C.C.A., von Itzstein M.;
RT "A conformational study of the human and rat encephalitogenic myelin
RT oligodendrocyte glycoprotein peptides 35-55";
RL Eur. J. Biochem. 246:59-70 (1997).
CC -1- FUNCTION: Minor component of the myelin sheath. May be involved in
CC completion and/or maintenance of the myelin sheath and in cell-
CC cell communication.
CC -1- SUBUNIT: May form homodimers.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
CC localized on the surface of myelin and oligodendrocyte cytoplasmic
CC membranes.
CC -1- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
CC ACTIVE MYELINATION.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
CC family.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
CC with oligodendrocyte-myelin glycoprotein (OMG).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M99485; AAA41628.1; --
DR EMBL; L21995; AAF74786.1; --
DR PIR; B47712; B47712.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IG; 1.

DR PROSITE; P50835; IG_LIKE; 1.
 KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
 FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 156 176 POTENTIAL.
 FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 209 223 POTENTIAL.
 FT DOMAIN 230 225 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 30 139 IG-LIKE.
 FT DISULFID 51 125 POTENTIAL.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC... (POTENTIAL).
 SQ SEQUENCE 245 AA; 27881 MW; C97F8AD6D6A32B4 CRC64;
 Query Match 7.5%; Score 77.5; DB 1; Length 245;
 Best Local Similarity 24.9%; Pred. No. 2.5; Indels 77; Gaps 10;
 Matches 49; Conservative 18; Mismatches 53;
 QY 1 MATLWGGLRLGSLLSCLALSLLVLLLAQLSDA-AKNE-----EDVRCKCI 45
 DB 1 MAGWNS-----LSLPSCLLSLLLLQLSRSVAGQFRVIGPHPIRALVGDDELPCR 52
 QY 46 CPYKENS-----HYNNKISQKDCCLHVPMP----- 76
 DB 53 ISPGKNTMEVGVYRSPFRVHLVRNG---KDQD-----AEQAPYRGRTLLKESIGE 105
 QY 77 -----VRGPDVEAY-CLRCCEKYERSSTVTKVITIIY-----LSILGL--LLLY 118
 DB 106 GKVALRQTVRFSDGEGYTCFFRDHSQVEEAELKVEDFPYINFGVLLALVLPVLL 165
 QY 119 MVTTLVEPIKRLFG 135
 DB 166 QVSVGLVFLQLRLG 182
 RESULT 13
 NOLIM HIPAM
 ID NOLIM HIPAM STANDARD; PRT; 318 AA.
 AC Q9ZZZ1;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 GN MTND1 OR ND1 OR NADH1.
 OS Hippopotamus amphibius (Hippopotamus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
 OX NCBI_TaxID=9833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99098146; PubMed=9881471;
 RA Ursing B.M., Arnason U.;
 RT "Analyses of mitochondrial genomes strongly support a hippopotamus-whale clade."
 RL Proc. R. Soc. Lond., B, Biol. Sci. 265:2251-2255(1998).
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -!- SIMILARITY: Belongs to the complex I subunit 1 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ010957; CAA09428.1; --
 DR InterPro; IPR001694; Resp_NADH_dh1.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 KW Ubiquinone; Mitochondrion; Oxidoreductase; NAD; Transmembrane.

SQ SEQUENCE 318 AA; 35928 MW; 2B63BBF6E3125B03 CRC64;
 Query Match 7.5%; Score 77.5; DB 1; Length 318;
 Best Local Similarity 38.9%; Pred. No. 3.3; Indels 7; Gaps 2;
 Matches 21; Conservative 12; Mismatches 14;
 QY 105 IITVLSLGLLLMYVLTVEPIKRLRFLFGHAQLIQSDDDDIGH---QPFANA 155
 DB 3 IINTLMVAPILLAMAFLLTVE---RKILGYMQLRKGNVIGPYGLLOPFADA 52
 RESULT 14
 ERB2_RAT
 ID ERB2_RAT STANDARD; PRT; 1257 AA.
 AC P06434;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor receptor-related protein).
 DE ERBB2 OR NEU.
 GN Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Neuroblastoma;
 RX MEDLINE=86113662; PubMed=3945311;
 RA Bargmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related protein."
 RL Nature 319:226-230(1986).
 RN [2]
 RP SEQUENCE OF 852-905 FROM N.A.
 RX TISSUE=Sciatic nerve;
 RX MEDLINE=91222560; PubMed=2025425;
 RA Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system."
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=92155181; PubMed=1346763;
 RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D., Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-oncogenic and oncogenic forms of the neu protein."
 RL EMBO J. 11:43-48(1992).
 CC -!- FUNCTION: Essential component of a neurotrophin-receptor complex, although neurotrophins do not interact with it alone. GP30 is a potential ligand for this receptor. Not activated by EGF, TGF-alpha and amphiregulin.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors. The constitutively activated oncogenic variant forms a homodimer. Interacts with PKCABP (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine residues (By similarity).
 CC -!- SIMILARITY: Belongs to the EGF receptor family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X03362; CAA27059.1; ALT_INIT.

DR PIR; A24562; TVRINU.
DR PDB; 1IIJ; 27-JUN-01.
DR DB; IN8Y; 18-FEB-03.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_Fac_recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP_2.
DR PRINTS; P00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 4.
DR SMART; SM00219; TyRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; MultiGene family; Receptor; Signal;
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Proto-oncogene; Disease mutation; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 1 21
FT DOMAIN 22 1257
FT TRANSMEM 22 654
FT DOMAIN 655 677
FT TRANSMEM 678 1257
FT DOMAIN 139 339
FT DOMAIN 473 646
FT DOMAIN 722 989
FT NP_BIND 728 736
FT BINDING 755 755
FT ACT_SITE 847 847
FT DISULFID 196 205
FT DISULFID 200 213
FT DISULFID 221 228
FT DISULFID 225 236
FT DISULFID 237 245
FT DISULFID 241 253
FT DISULFID 256 265
FT DISULFID 269 296
FT DISULFID 300 312
FT DISULFID 316 332
FT DISULFID 335 339
FT DISULFID 513 522
FT DISULFID 517 530
FT DISULFID 533 542
FT DISULFID 546 562
FT DISULFID 565 578
FT DISULFID 569 586
FT DISULFID 589 598
FT DISULFID 602 625
FT DISULFID 628 636
FT DISULFID 632 644
FT MOD_RES 1141 1141
FT MOD_RES 1250 1250
FT CARBOHYD 68 68
FT CARBOHYD 188 188
FT CARBOHYD 260 260
FT CARBOHYD 532 532
FT CARBOHYD 573 573
FT CARBOHYD 631 631
FT VARIANT 661 661
SQ SEQUENCE 1257 AA; 61.29264583011402 CRC64;
Query Match 7.5%; Score 77.5; DB 1; Length 1257;
Best Local Similarity 27.4%; Fred. No. 15;
Matches 29; Conservative 19; Mismatches 25; Indels 33; Gaps 7;
QY 46 CPYKNSGHIYKNKISQKDCDLHVPE-----MPV-RGPDEAYCLRCE----- 90

RESULT 15

MUSE_HUMAN
ID MUSEB_HUMAN STANDARD; PRT: 5703 AA
AC Q9HC84; Q00447; Q00573; Q14985; Q15494; Q95291; Q95451; Q14881;
AC Q99552; Q9UE28;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High
DE molecular weight salivary mucin M61) (Sublingual gland mucin).
GN MUC5B OR MUC5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-1594 FROM N.A.
RA Chen Y.; Di Y.P.; Wu R.;
RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
RT human MUC5B mucin gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-1325 FROM N.A.
RA Ofner G.D.; Nunes D.P.; Keates A.C.; Afshal N.H.; Troxler R.F.;
RX MEDLINE=99009274; PubMed=9790959;
RT "The amino-terminal sequence of MUC5B contains conserved
RT multifunctional D domains: implications for tissue-specific mucin
RT functions.";
RL Biochem. Biophys. Res. Commun. 251:350-355(1998).
RN [3]
RP SEQUENCE OF 40-1324 FROM N.A.
RX MEDLINE=99023932; PubMed=9804771;
RA Desseyn J.-L.; Eulsine M.P.; Porchet N.; Aubert J.-P.; Laine A.;
RT "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
RT sequences upstream of the large central exon.";
RL J. Biol. Chem. 273:30157-30164(1998).
RN [4]
RP SEQUENCE OF 1326-4895 FROM N.A.
RX TISSUE=Placenta;
RA Desseyn J.-L.; Guyonnet-Duperat V.; Porchet N.; Aubert J.-P.;
Laine A.;
RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
RT various alternate subdomains resulting in a super-repeat. Structural
RT evidence for a 11p15.5 gene family.";
RL J. Biol. Chem. 272:3168-3178(1997).
RN [5]
RP SEQUENCE OF 4057-4480 FROM N.A.
RX TISSUE=Salivary gland;
RA MEDLINE=97292540; PubMed=9147051;
Hannibal J.; Clausen H.;
RT "Identification of a major human high molecular weight salivary mucin
RT (M61) as tracheobronchial mucin MUC5B.";
RL Glycobiology 7:413-419(1997).
RN [6]
RP SEQUENCE OF 4721-5703 FROM N.A.
RX TISSUE=Gall bladder;
RA MEDLINE=97293229; PubMed=9164870;
Keates A.C.; Nunes D.P.; Afshal N.H.; Troxler R.F.; Offner G.D.;
RT "Molecular cloning of a major human gall bladder mucin: complete C-
RT terminal sequence and genomic organization of MUC5B.";
RL Biochem. J. 324:295-303(1997).
RN [7]
RP SEQUENCE OF 4809-5687 FROM N.A.

RC TISSUE=Sublingual gland;
RX MEDLINE=96125355; PubMed=8554565;
RA Troxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim F.G.;
RT "Molecular cloning of a novel high molecular weight mucin (MGI)
from human sublingual gland";
RL Biochem. Biophys. Res. Commun. 217:1112-1119 (1995).
RN [8]
RP SEQUENCE OF 4859-5703 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97347489; PubMed=9201995;
RA Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;
RT "Genomic organization of the 3 region of the human MUC5B mucin.";
RL J. Biol. Chem. 272:16873-16883 (1997).
CC -!- FUNCTION: Salivary mucin that is thought to contribute to the
lubricating and viscoelastic properties of whole saliva.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
in submaxillary glands, endocervix, gall bladder, and pancreas.
CC -!- PM: Highly glycosylated.
CC -!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
CC -!- SIMILARITY: Contains 3 WFCD domains.
CC -!- SIMILARITY: Contains 4 WFCD domains.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF107890; AAC33673.1; -;
DR EMBL; AF086604; AAC67545.1; -;
DR EMBL; AJ004862; CA06167.1; -;
DR EMBL; Z7496; CA95577.1; -;
DR EMBL; X74955; CA52910.1; -;
DR EMBL; U63836; AAB61398.1; -;
DR EMBL; U78554; AAC51344.1; -;
DR EMBL; U78552; AAC51344.1; JOINED.
DR EMBL; U78553; AAC51344.1; JOINED.
DR EMBL; U78551; AAC51343.1; -;
DR EMBL; U93031; AAB65151.1; -;
DR EMBL; Y09788; CAA70926.1; -;
DR EMBL; X00770; -;
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR009041; PMP_inhibitor.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR006552; VC_Out.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot_1.
DR Pfam; PF01826; TIL_1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00094; vwd; 4.
DR SMART; SM00214; VWC; 6.
DR SMART; SM00215; VWC_Out; 4.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWF_1; 2.
DR PROSITE; PS01208; VWF_2; 2.
DR PROSITE; PS01208; VWF_3; 2.
KW Glycoprotein; Repeat; Signal; Polymorphism.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 5703 MUCIN 5B.
FT DOMAIN 77 225 VWF 1.
FT DOMAIN 329 386 TIL.
FT DOMAIN 426 580 VWF 2.
FT DOMAIN 658 918 VWF 1.
FT DOMAIN 896 1044 VWF 3.

FT DOMAIN 1457 1603 THR-RICH.
FT DOMAIN 1609 1609 THR-RICH.
FT DOMAIN 5005 5178 VWF 4.
FT DOMAIN 5353 5425 VWF 2.
FT DOMAIN 5462 5528 VWF 3.
FT DOMAIN 5594 5683 CTCK.
FT DISULFID 5594 5646 BY SIMILARITY.
FT DISULFID 5622 5660 BY SIMILARITY.
FT DISULFID 5626 5676 BY SIMILARITY.
FT DISULFID 5643 5678 BY SIMILARITY.
FT DISULFID 5645 5682 BY SIMILARITY.
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 516 516 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 806 806 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 930 930 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1293 1293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1557 1557 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1775 1775 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2192 2192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2721 2721 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3419 3419 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3948 3948 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4745 4745 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4901 4901 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4958 4958 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4965 4965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4987 4987 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5037 5037 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5052 5052 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5156 5156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5427 5427 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5467 5467 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5506 5506 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5507 5507 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5543 5543 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5553 5553 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5604 5604 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5618 5618 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5662 5662 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 5137 5137 T -> S (in dbSNP:2672788).
FT CONFLICT 34 34 /FTid=VAR_014123.
FT CONFLICT 95 100 G -> E (IN REF. 2).
FT CONFLICT 104 104 PFGLCN -> LPCLCK (IN REF. 2).
FT CONFLICT 142 104 S -> C (IN REF. 2).
FT CONFLICT 142 142 E -> K (IN REF. 1).
FT CONFLICT 225 225 R -> S (IN REF. 2).
FT CONFLICT 330 331 PL -> T (IN REF. 2).
FT CONFLICT 337 337 E -> N (IN REF. 2).
FT CONFLICT 356 356 E -> K (IN REF. 2).
FT CONFLICT 362 362 G -> R (IN REF. 2).
FT CONFLICT 369 369 MISSING (IN REF. 2 AND 3).
FT CONFLICT 374 374 D -> N (IN REF. 2).
FT CONFLICT 393 394 RT -> TR (IN REF. 2).
FT CONFLICT 468 469 RK -> GR (IN REF. 2).
FT CONFLICT 512 512 L -> P (IN REF. 2).
FT CONFLICT 585 587 GAA -> AH (IN REF. 3).
FT CONFLICT 601 601 A -> S (IN REF. 3).
FT CONFLICT 628 629 DP -> RS (IN REF. 2).
FT CONFLICT 633 633 F -> L (IN REF. 2).
FT CONFLICT 676 676 A -> P (IN REF. 3).

Query Match 7.5%; Score 77.5; DB 1; Length 5703;

Best Local Similarity 28.2%; Pred. No. 80;

Matches 31; Conservative 11; Mismatches 37; Indels 31; Gaps 7;

QY 6 GGLRLIGSL--SLSLALSVLLAQLSDAAK-----NFEDVRC-----KCICP 47

Db 780 GKLSCLGASLQKSTGCAAPMVYLDGNSNSAGTPGACLRSLCHTLDVGCFSTHCVSGCVP 839

Mon Mar 15 08:35:15 2004

QY 48 P--YKENSHTYNKNISQDCDCLH---VVEPMFVRGPDVEAYCLRCECK 92
| : | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 840 PGLVDSGGGC----IAEEDPCVNNEATYKP-----GETIRVDCNTCTCR 881

Search completed: March 15, 2004, 08:26:15
Job time : 20 secs

Query Match	76.0%;	Score 784;	DB 13;	Length 196;
Best Local Similarity	79.8%;	Pred. No. 8.5e-78;		
Matches 148;	Conservative 18;	Mismatches 20;	Indels 0;	Gaps 0;
13	SLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSIGHYNNKNSQKDCDCLHHV	72		
11	NLLTFSSLCLFLLLSAQTSEAKNSDIRCKICPPYKEVDGKIYQNVSLKDCNCLHHV	70		
73	EPMPVPGDVAYCLRCCKCYEERSVTKVITIIYLSIIGLLIYVYVLTVEPIIKR	132		
71	EPVPEGVGQVAYCLRCCKCYEERSSGTKCTIIYLSIIGLLIYVLTVEPIIKR	130		
133	LFHGAQLIQSDDDIGDHQPPANAHDLVLAISRSPANVLNKVEYAQQRWKLQVQQRKSVD	192		
131	LFHGSQLIQSDDDVGDDQPPANAHDLVLSRSPANVLNKVEHAQQRWRRQVQQRKSVD	190		
193	RHVVL	198		

OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Cyprinidae;
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045338; AHA45338.1; -
DR InterPro; IPR004153; CXKC-repeat.
DR Pfam; PF03128; CXKC; 1.
KW Hypothetical protein.
SQ SEQUENCE 186 AA; 21430 MW; C334B4326C605279 CRC64;
Query Match 52.7%; Score 543.5; DB 13; Length 186;
Best Local Similarity 59.7%; Pred No. 2e-51;
Matches 108; Conservative 29; Mismatches 33; Indels 11; Gaps 4;
QY 20 LALSLL-LAQLSDAAKFNEDVRCIKCPYKENGSHYKNISOKDCDCLHVPMPVR 78
DB 15 LAPVLIEVVSFASADKSPEDVRCKICPYPYRNITGHYKQNTQKDCNCLHVPMPVR 74
QY 79 GPVVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYMYVLTIVPEILKRLFGHAQ 138
DB 75 GHVVEAYCLRCCKYEERSVTKVTIIYLSVVGALLYMLFLLLIDPLI-RKHPYTM 133
QY 139 LIQSDDDDDIGHQPFANAHVLAERSRAN-VLNKVEYAQORWKLQVQEQKSVFDRH 197
DB 134 PLQNEEDSEVVRPRVDG-----ARGNTVLERVEGAQORWKLQVQEQKSVFDRH 185
QY 198 S 198
DB 186 S 186
RESULT 4
Q21960 PRELIMINARY; PRT; 261 AA.
ID Q21960
AC Q21960;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN R12C12.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RT None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Favello A.;
RT "The sequence of C. elegans cosmid R12C12.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U23510; AAC46785.1; -
DR PIR; T16732; T16732.
DR WormPep; R12C12.6; CE02851.

DB 191 RHVLS 196
RESULT 2
Q72649 PRELIMINARY; PRT; 124 AA.
ID Q72649
AC Q72649;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Clorif5 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shemen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buecaw K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.P.; Casavant T.L.; Scheetz T.B.;
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
RA Bosak S.A.; McSwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalon D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fahey J.; Helton E.; Kettman M.; Madan A.; Rodrigues S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.;
RA Krzywinski M.I.; Skalska U.; Smaluk D.E.; Schnerch A.; Schein J.E.;
RA Jones S.J.; Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015884; AHA15884.1; -
SQ SEQUENCE 124 AA; 14428 MW; 1D0CC22276D6E62 CRC64;
Query Match 61.6%; Score 635; DB 4; Length 124;
Best Local Similarity 100.0%; Pred No. 1.2e-61;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 75 MPVGPVVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYMYVLTIVPEILKRLR 134
DB 1 MPVGPVVEAYCLRCCKYEERSVTKVTIIYLSILGLLLYMYVLTIVPEILKRLR 60
QY 135 GHQLIQSDDDIGHQPFANAHVLAERSRANVLNKNVEYAQORWKLQVQEQKSVFDRH 194
DB 61 GHQLIQSDDDIGHQPFANAHVLAERSRANVLNKNVEYAQORWKLQVQEQKSVFDRH 120
QY 195 VWLS 198
DB 121 VWLS 124
RESULT 3
Q72W08 PRELIMINARY; PRT; 186 AA.
ID Q72W08
AC Q72W08;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.

```
DR InterPro: IPR008853; TMEM9.
DR Pfam: PF05434; TMEM9; 1.
KW Hypothetical protein.
SQ SEQUENCE 261 AA; 29923 MW; 278F2D3611A06964 CRC64;

Query Match      26.9%; Score 277.5; DB 5; Length 261;
Best Local Similarity 32.1%; Pred. No. 4.8e-22;
Matches 68; Conservative 44; Mismatches 65; Indels 35; Gaps 8;

QY 20 LALSULLAQLSDAA--KNFVRCCKICP-----PKYKNS-----CHIYNK-NIS 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 52 LFIALFLPALSAQAGTEANFETRCRCICPSLLKFLDLAENTTEKTEGLRRFYTKNIE 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 63 QKDCDCLHVPEMPVRGPD---VEAYCLRCCKEYBERSVTKIITIIYLSIGLILLLYM 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 112 PSCKEPSNIKQOVSNFVDETHWDAFLANDCRYESRNTVLKVVIFVICVIAVLGYM 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 120 VYLTIVPEPIL-KRRLFGHAQLIQSDDDIGDH-----QPFANAHDVLARSRA 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 172 VFMLCLDPLMKRLKLSISYQ--QHNDMEDNIFAAAPSTDDESSASNSMDTQGTTRARS 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 167 NVLNKVEYAQQRWKLQVQQRKSVFDRHVLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 230 NVLGRVEAQNRWKKVVEQRNIFEDHTMLN 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
Q8SXE9 PRELIMINARY; PRT; 227 AA.
AC Q8SXE9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RH45888p.
GN CG1161.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreesen D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cealnikar S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY089673; AAL90411.1; -.
DR FlyBase; FBgn037313; CG1161.
DR InterPro; IPR008853; TMEM9.
DR Pfam; PF05434; TMEM9; 1.
SQ SEQUENCE 227 AA; 24649 MW; 9E7040F3BCA34B49 CRC64;

Query Match      22.0%; Score 226.5; DB 5; Length 227;
Best Local Similarity 31.6%; Pred. No. 1.6e-16;
Matches 50; Conservative 38; Mismatches 63; Indels 7; Gaps 4;

QY 43 KTCIPPKYKNSGHYNNKISQKDCDCLHVPEP-MFVRGDDVEAYCLRCCKEYERSVTI 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 75 KVIGQPVLAPAGKNSSSSTTECVACAGALLPRDLANGKEL-PICAEKCKSHVARTILI 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 102 KVTIIIVLSIGLILLVWYLTIVPEPILKRL-FGHAQLIQSDDDIGDHPFANAHDVLA 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 KVVIVIVISILVIMFLMCLDPLNKRKVKANYQEHTNEDETPPLPAVNNQELSA 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 161 RRSRANLVNKVEYAQQRWKLQVQQRKSVFDRHVLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 194 ----RANVLNRVGHQDQKWKQVRQEQRRHYDRTHTMLN 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
Q86FF8 PRELIMINARY; PRT; 210 AA.
AC Q86FF8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Clone ZD1129 mRNA sequence.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu F., Hu W., Yan Q., Xu X., Zhu Z., Zhang X., Wang J., Sun J.,
RA Xu X., Wang Z., Zeng L., Rong Y., Wu X., Qu J., Xu Z., Huang J.,
RA Ma Y., Wang S., Wang Z., Xue C., Peng Z., Chen Z., Han Z.;
RT "The full-length cDNA of S. japonicum genes.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY222905; AAF05917.1; -.
DR EMBL; AY222905; AAF05917.1; -.
SQ SEQUENCE 210 AA; 23447 MW; 665FD670BB29FDB0 CRC64;

Query Match      19.1%; Score 197; DB 5; Length 210;
Best Local Similarity 33.1%; Pred. No. 2.6e-13;
Matches 46; Conservative 27; Mismatches 46; Indels 20; Gaps 5;

QY 15 LSLSLALSULLAQLSDAAKQFEDVRCYCIC---PPYKEN---SCHYNNKISQKDCDC 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MKLFSLLIGAFILHTASAA---YEDARCKVCCLPGSYMNTKSKKIYVRAIPSEKIC 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 69 LHVPEMPVRGPDVEAYCLRCCKEYERSVTKIITIIYLSIGLILLVWYLTIVPEI 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 EYMLQ-----EDKELCHFECKFQVRNTTIKVVVCFLITITAILYMLFLLLEPL 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 129 LKRLFGHAQLIQSDDDIG 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 111 LSTRLSHKTMI----DLG 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
Q812L1 PRELIMINARY; PRT; 1782 AA.
AC Q812L1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN PF11470C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Omond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL929358; CAD51980.1; -.
DR InterPro; IPR001611; LRR.
```

```
DR InterPro; IPR003885; LRR_CYST.
DR InterPro; IPR007092; LRR_SDS22.
DR Pfam; PF00560; LRR; 5.
DR SMART; SMO0365; LRR_SD22; 5.
KW Hypothetical protein.
SQ SEQUENCE 1782 AA; 214024 MW; 8DD60393FDC2CB50 CRC64;

Query Match      8.8%; Score 91; DB 5; Length 1782;
Best Local Similarity 21.0%; Pred. No. 1.2;
Matches 49; Conservative 36; Mismatches 64; Indels 84; Gaps 13;

QY 32 DAAKNFEDVRCKICP-PYKENSCHIYNK-----NISKQDCDCLHV 72
DB DHPREFENMMK-KCPNDYKNI-EIKNKEQINNSVIKRIEATSKSLNIGKDDLLNIV 444
QY 73 EPM--PVRGPPVEAYCLRCEKYEERSV-TIKVTI-----LIY 113
DB EDKEEFTYTTDVTNCTNNNNCIYIDKEKYNKIKYIENNIKFNPFKDLINELMNIWS 504
QY 114 LLLLYMYVLTVEPIKRE-----RLFG-----HAQIQSD 143
DB LNNLIEIPKTIKSKIKKINDNITDYFKGLDELEENFTQILNFFSEVKNNDHLSIVLSE 564
QY 144 DDIGHQHFANAHVLAERSRANVLNKE-YAQRW-----KLVQVQRKSVF 191
DB DEINEYYSYKN-----YRSNMNMLEYISSSYNKAGSILIEKKHLF 608

RESULT 8
Q9B9D8 PRELIMINARY; PRT; 119 AA.
AC Q9B9D8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
oxidoreductase chain 1) (Fragment).
GN NDI.
OS Spodoptera littoralis (Egyptian cotton leafworm).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7109;
RN [1]
RP SEQUENCE FROM N.A.
RA Abraham D., Ryrholm N., Wittzell H., Scoble M.J., Holloway J.D.,
RA Lofstedt C.;
RT "Molecular phylogeny of the subfamilies in Geometridae (Geometroidea:
RT Lepidoptera)".
RL Mol. Phylogenet. Evol. 0:0-0(2001).
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
DR EMBL; AF178857; AAK00984.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001694; Resp_NADH_dhl.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
FT NON_TER 119
SQ SEQUENCE 119 AA; 13739 MW; 3AC1FA52853252C3 CRC64;

Query Match      8.2%; Score 84.5; DB 8; Length 119;
Best Local Similarity 29.9%; Pred. No. 0.31;
Matches 23; Conservative 18; Mismatches 23; Indels 13; Gaps 3;

QY 107 IYLSILGLLLY-----MVYLTVEPIKRLRFGHAQLIQSDDDIGDH---QPFANAH 157
DB IYLSILGLLLY-----MVYLTVEPIKRLRFGHAQLIQSDDDIGDH---QPFANAH 157
QY 158 VLARSRRANVLNKEVY 174
DB LFTKEQTYPNFSNLSY 76
60 LFTKEQTYPNFSNLSY 76

RESULT 10
Q8WRC9 PRELIMINARY; PRT; 1453 AA.
AC Q8WRC9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Cysteine repeat modular protein 4 PDCRM4 (Fragment).
OS Plasmodium berghei (strain Anka).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5823;

Query Match      8.2%; Score 84.5; DB 8; Length 235;
Best Local Similarity 29.9%; Pred. No. 0.66;
Matches 23; Conservative 18; Mismatches 23; Indels 13; Gaps 3;

QY 107 IYLSILGLLLY-----MVYLTVEPIKRLRFGHAQLIQSDDDIGDH---QPFANAH 157
DB IYLSILGLLLY-----MVYLTVEPIKRLRFGHAQLIQSDDDIGDH---QPFANAH 157
QY 158 VLARSRRANVLNKEVY 174
DB LFTKEQTYPNFSNLSY 76
60 LFTKEQTYPNFSNLSY 76
```



```
DE OSJNB0059K02.19 protein.
GN OSJNB0059K02.19.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;
[1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Yu Z., Chen L., Fan D.L., Wang Q.J., Zhang L., Li T.,
RA Lu Y.Q., Mu S.L., Liu X.H., Lu T.F., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu P.X., Qian Y.M., Yang K., Zhou B., Chen Z.H.,
RA Hao P., Zhang H., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding G.F.,
RA Gu J.L., Chen S.T., Ni L., Zhu P.H., Hong G.F.,
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL06692; CAB04509.1; 81BCF310A35088F1 CRC64;
SQ SEQUENCE 783 AA; 86458 MW; 81BCF310A35088F1 CRC64;

Query Match 8.1%; Score 83.5; DB 10; Length 783;
Best Local Similarity 20.8%; Pred. No. 3.3;
Matches 47; Conservative 28; Mismatches 72; Indels 79; Gaps 10;

QY 5 WGLRLGLSLSLCLALSLLAQLSDAAKXNFDVRCCKICPPYKNSGHIYNNK-NISQ 63
DB 287 WLDLRDGAIFLS-----VGPNITDFRCLSL-----HSSYINDLNDR 326

QY 64 KDCDCLHVPEP-----VRGPDV---EAYCLRCCKYEERSSVTIK-----VT 104
DB 327 RCNCSQYEGNPIYRDGCRDDEQCPDVYFCHGTCTNMPGTVCRLAKSVKSLPGLIT 386

QY 105 ILYLSILGLLLLYMYLVFLVEPI-----LKRRLP--GHA-----QLIQSDDDIGDHQP 151
DB 387 IIAVSAGFGLLSLLGVAKITNKIKQRAKLRKFFKXNGLLLQLQLISSNKDIAERMK 446

QY 152 FANAHDVLAERSRANVLNKVEAQQRWKLVQVEQRKSVFDRHVL 197
DB 447 I-----FSLELDQATNFKDQNRIL 466

RESULT 14
Q869E1 PRELIMINARY; PRT; 1192 AA.
AC Q869E1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Xenopus laevis (African clawed frog) DNA ligase I (EC
DE 6.5.1.1) (Polydeoxyribonucleotide synthase).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OC NCBI_TaxID=44689;
[1]
RP SEQUENCE FROM N.A.
RA STRAIN=AX4;
RC MEDLINE=22092622; PubMed=12037910;
RA Gloeckner G., Eichinger L., Szafarski K., Pachebat J., Dear P.,
RA Lehman B., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer J.M., Rosenthal A., Noegel A.A.;
RA "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RA Nature 418:79-85(2002).
[2]
RP SEQUENCE FROM N.A.
RA STRAIN=AX4;
RC Baumgart C.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC116957; AA052526.1; -.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.

DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000977; DNA_ligase.
DR Pfam; PF01068; DNA_ligase; 1.
DR Pfam; PF04879; DNA_ligase_A_C; 1.
DR Pfam; PF04875; DNA_ligase_A_N; 1.
DR TIGRPFAM; TIGR00574; dnll; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR PROSITE; PS00333; DNA_LIGASE_A2; 1.
DR PROSITE; PS0160; DNA_LIGASE_A3; 1.
KW Ligase.
SQ SEQUENCE 1192 AA; 137513 MW; B52A5842CD779766 CRC64;

Query Match 8.1%; Score 83; DB 5; Length 1192;
Best Local Similarity 23.6%; Pred. No. 5.9;
Matches 46; Conservative 28; Mismatches 59; Indels 62; Gaps 11;

QY 11 LGSLSLSLCLALS---LLLAQ---LSDAAKXNFDVRCCKICPPYK-----E 51
DB 801 IDNLSCTSLKVGIPVKPMLAQPTTGISQMLDRFSDMEFTC---EFKYDGERAQIHLRPD 857

QY 52 NSGHIYNNKI-----SQKDCDCLHVPEPVRGPDVAYCLRCSC-----K 92
DB 858 GTTHIYTRNLEDYTKYQYDI--VANVTKFVGNVKSFLDCEAVAFDAATKKILSFQVLS 915

QY 93 YEERSV---TKVTIIILYLSILGLLLMYVLTIV-EPILKRRLFQHAQLIQSD----- 143
DB 916 TRARKSVQLSQIKVPKVCVF---APLLYLNGSLIDEPLIKRREHLVENFIASDGVFAF 971

QY 144 -----DDIGDHQP 152
DB 972 AKYSNISVDNIOQSY 986

RESULT 15
Q9B9D9 PRELIMINARY; PRT; 120 AA.
AC Q9B9D9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
DE oxidoreductase chain 1) (Fragment).
GN NDI.
OS Agrotis segetum (turnip moth).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Noctuinae; Agrotis.
OX NCBI_TaxID=47767;
[1]
RP SEQUENCE FROM N.A.
RA Abraham D., Ryrholm N., Wittzell H., Scoble M.J., Holloway J.D.,
RA Lofstedt C.;
RT "Molecular phylogeny of the subfamilies in Geometridae (Geometroidea:
RT Lepidoptera)";
RL Mol. Phylogenet. Evol. 0:0-0(2001).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
DR EMBL; AF178856; AA000983.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016431; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001694; Resp_NADH_dhl.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_NDI_1; 1.
KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
FT NON_TER 1
FT NON_TER 120
FT NON_TER 120
SQ SEQUENCE 120 AA; 13829 MW; 2295E38A68675A0E CRC64;
```

2
2
4
.
5
4
4

7
-
7

-
-
7

2
5

Query Match 8.0%; Score 82.5; DB 8; Length 120;

Search completed: March 15, 2004, 08:27:14

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:23:48 ; Search time 21 Seconds
(without alignments)
906.949 Million cell updates/sec

Title: US-09-997-573-418
Perfect score: 1031
Sequence: 1 MATLWGLRLGSLSLSLCL.....WKLQVQRKSVFDRHVLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530	51.4	187	T30536	hypothetical prote
2	277.5	26.9	261	T16732	hypothetical prote
3	84.5	8.2	235	A4233	NADH2 dehydrogenas
4	80.5	7.8	432	T37684	probable 6-phospho
5	79	7.7	479	D86340	protein F2D10.30 [
6	78.5	7.6	450	AF0207	probable membrane
7	78.5	7.6	841	T01011	hypothetical prote
8	78.5	7.6	1700	S08167	Balbani ring 3 pr
9	78	7.6	285	S73465	MG011 homolog D12
10	77.5	7.5	275	S75818	hypothetical prote
11	76.5	7.4	2471	T42977	large tegument pro
12	75.5	7.3	315	QXMSIM	NADH2 dehydrogenas
13	75.5	7.3	401	A3537	oligopeptide trans
14	75.5	7.3	402	S46641	probable membrane
15	75.5	7.3	617	S48160	metalloproteinase
16	75	7.3	297	T08589	hypothetical prote
17	75	7.3	325	T47229	arginase (EC 3.5.3
18	75	7.3	358	T47228	hypothetical prote
19	74.5	7.2	318	T11506	NADH2 dehydrogenas
20	74.5	7.2	1372	T25933	hypothetical prote
21	74.5	7.2	1630	S64403	ESPI protein - yea
22	74	7.2	246	A47712	myelin/Oligodendro
23	74	7.2	256	B32393	T-cell antigen 4-1
24	74	7.2	606	JC5604	ABC-transporting p
25	74	7.2	1003	1 PYVZAM	spheroidin precurs
26	74	7.2	1304	1 A46546	leukocyte common a
27	73.5	7.1	1260	1 TVRTNU	protein-tyrosine k
28	73.5	7.1	1958	1 T39808	hypothetical prote
29	73	7.1	527	2 T23699	hypothetical prote

30 73 7.1 599 2 T25835 hypothetical prote
31 73 7.1 896 2 I45858 desmocollin - bovi
32 72.5 7.0 318 2 A58850 NADH2 dehydrogenas
33 72.5 7.0 318 2 S41820 NADH2 dehydrogenas
34 72.5 7.0 350 2 C83608 probable transcrip
35 72.5 7.0 351 2 AC2289 hypothetical prote
36 72.5 7.0 735 2 T00850 probable receptor-
37 72 7.0 584 2 S76210 hypothetical prote
38 72 7.0 1207 1 EGHU epidermal growth f
39 72 7.0 1582 2 T15308 hypothetical prote
40 72 7.0 5762 2 A41819 proline-rich pepti
41 71.5 6.9 323 2 A34284 NADH2 dehydrogenas
42 71.5 6.9 1264 2 S41603 type V adenylyl cy
43 71.5 6.9 1321 2 JE0352 mucin MUC5B, trach
44 71 6.9 481 2 AE1207 two-component sens
45 71 6.9 940 2 T31575 hypothetical prote

ALIGNMENTS

RESULT 1

T30536
hypothetical protein 2 - Fugu rubripes (fragment)

C:Species: Fugu rubripes

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T30536

R:Riboldi Tunncliffe, G.R.; Platzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Ros
submitted to the EMBL Data Library, September 1997

A:Description: Analysis of the genomic loci of Fugu rubripes homologs of the human dis
A:Reference number: Z20848

A:Accession: T30536

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-187 <R18>

A:Cross-references: EMBL:AF026198; NID:G3098263; PID:G3098268; PID:AACI5584.1

C:Genetics:

A:Note: PUT2

A:Note: Intron positions not resolved (incomplete sequence)

Query Match 51.4%; Score 530; DB 2; Length 187;

Best Local Similarity 57.8%; Pred. No. 3.7e-43;

Matches 108; Conservative 29; Mismatches 40; Indels 10; Gaps 4;

QY 14 LLSLSCLSLVLLAQLSDA--AKNFEDVRCKCIPPYKENSGLHYNNKISQKQDCCLHV 71
Db 8 LWMALAFALMTLFLDNGVGTQAKSFDDVRCKCIPPYRNISGLHYNNFTQKNCNCLHV 67
QY 72 VBPMPVRGPDVEAYCLRCCKEYERSSVTIKVTIIYLSILGLLLYVMVLTLPILKR 131
Db 68 VDPMPVPGNDVEAYCLLCKEYERSTWRTVIIIFLSVVGALLLYMLFLLDPLI-R 126
QY 132 RLFGHAQLIQSDDDIGDHQPPANAHDLVARSRRAN-VLNKVEYAOQKVKQVQQRKSV 190
Db 127 KPDPPLAQLLNHEEDSEDIQP-----QMSGDPARGNTVLERVEGAQQRWKQVQQRKTIV 180
QY 191 FDRHVVL 197
Db 181 FDRHKML 187

RESULT 2

T16732
hypothetical protein R12C12.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999

C:Accession: T16732

R:Favella, T.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid R12C12.

A:Reference number: Z18568

A:Accession: T16732

A>Status: preliminary; translated from GB/EMBL/DBJ

submitted to the EMBL Data Library, October 1999
A:Reference number: Z21736
A:Accession: T37684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-432 <MCD>
A:Cross-references: EMBL:AL132675; PIDN:CAB59697.1; GSPDB:GN00066; SPDB:SPAC144.17C
A:Experimental source: strain 972h; cosmid c144
C:Genetics:
A:Gene: SPDB:SPAC144.17C
A:Map position: 1
A:Introns: 107/1
C:Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phos
F:259-430/Domain: phosphoglycerate mutase homology <PGMH>

Query Match 26.9%; Score 27.5; DB 2; Length 261;
Best Local Similarity 32.1%; Pred. No. 6.2e-19;
Matches 68; Conservative 44; Mismatches 55; Indels 35; Gaps 8;

QY 20 LALSIVLLAQLSDAA--KNFEDVRCKICP-----PYKNS-----GHYNNK-NIS 62
DB 52 LFIALLFALPALSQAGTEANFEDTRCICPCLLKFLDLAENTTEKTEGLRRRFTKTNIE 111
QY 63 QKDCDLHVPEMPVRGPD---VAYCLRCCKEYESSVTIKVTIILYLSILGLLLYM 119
DB 112 PSHKPSNIVKQVSNFVDETHMDFLANCDRIESNVLKVVIFVICVIAVLGYM 171
QY 120 VYLTVLVEPIL-KRRLFGLHAQLIQSDDDIGDH-----OPFANAHDVLARSRSRA 166
DB 172 VFLMCLDPLMKRLISYQ--QHNDMEDNIFAAAPSTDDSSASNSMDTQGTTRS 229
QY 167 NVLKNVEYAQRWKLQVQKQSVFDRHVLVS 198
DB 230 NVLGRVEAFQNRWKKVBEQRNIPFEDHTMLN 261

RESULT 3
A44233
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - fall armyworm mitochondrion (fra
N:Alternate names: NADH-ubiquinone oxidoreductase chain 1
C:Species: mitochondrion Spodoptera frugiperda (fall armyworm)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: A44233
R:Pasley, D.P.; Ke, L.D.
Mol. Biol. Evol. 9, 1061-1075, 1992
A:Title: Sequence evolution in mitochondrial ribosomal and ND-1 genes in lepidoptera: im
A:Reference number: A44233; MUID:193061985; PMID:1435234
A:Accession: A44233
A:Molecule type: DNA
A:Residues: 1-235 <PAS>
A:Cross-references: GB:M76713; NID:G43352; PIDN:AAA32079.1; PID:G552886
A:Note: sequence extracted from NCBI backbone (NCBIP:118938)
C:Genetics:
A:Gene: ND-1
A:Genome: mitochondrion
A:Start codon: ATA
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 1
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 8.2%; Score 84.5; DB 2; Length 235;
Best Local Similarity 29.9%; Pred. No. 1.4;
Matches 23; Conservative 18; Mismatches 23; Indels 13; Gaps 3;

QY 107 IYLSILGLLLY-----MVYLFVPEILKRLFGHAQLIQSDDDIGDH---OPTANAH 157
DB 4 IYMFLLGLLIIIGLVAVYLTLE---RKVLGYIQIRKGNKVGFGILQPFSDAIK 59

QY 158 VLARSRSRANVLNVEY 174
DB 60 LFTKEQTPNFSNLSY 76

RESULT 4
T37684
Probable 6-phosphofructo-2-kinase/fructose-2, 6-bisphosphate - fission yeast (Schizosa
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T37684
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.

submitted to the EMBL Data Library, October 1999
A:Reference number: Z21736
A:Accession: T37684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-432 <MCD>
A:Cross-references: EMBL:AL132675; PIDN:CAB59697.1; GSPDB:GN00066; SPDB:SPAC144.17C
A:Experimental source: strain 972h; cosmid c144
C:Genetics:
A:Gene: SPDB:SPAC144.17C
A:Map position: 1
A:Introns: 107/1
C:Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phos
F:259-430/Domain: phosphoglycerate mutase homology <PGMH>

Query Match 7.8%; Score 80.5; DB 2; Length 432;
Best Local Similarity 25.6%; Pred. No. 6.3;
Matches 45; Conservative 26; Mismatches 50; Indels 55; Gaps 11;

QY 19 CLA-LSVLLAQ-----LSDAANKN--FEDVRCKICPPYKENSQHI 56
DB 123 CLADLEIFLLKEGQVAYDATNGTRRRILYDRFKCGFKILFIESLC-----NKEDV 177
QY 57 YNNKISQKDCDLHVPEMPVRGPDVE---AYCLRCB---CKYE---ERSVTIKV--- 103
DB 178 INANIQ-----EAIHVSEEF--RNWOLEMAEKYCRIDILKCHYETIDEKDYSFVYMINF 231
QY 104 --TIIYLSILGLLLYVYLVPEILKRRLF-----GHAQLIQSDDDIGDHQ 150
DB 232 AETIANKSNEGYSLLSRILFLMNTLARKRFLVLPKASMRPLKREPEDIENRQ 287

RESULT 5
D86340
Protein F2D10.30 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D86340
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <STO>
A:Cross-references: GB:AE005172; NID:G8886943; PIDN:AAF80629.1; GSPDB:GN00141
C:Genetics:
A:Gene: F2D10.30
A:Map position: 1

Query Match 7.7%; Score 79; DB 2; Length 479;
Best Local Similarity 25.5%; Pred. No. 9.7;
Matches 35; Conservative 17; Mismatches 47; Indels 38; Gaps 6;

QY 14 LLSLSCLALSIVLLAQLSDAAKNFEDVRCKICPPYKEN-----SCHYNNKISQKDCD 67
DB 13 LLKLDPKSLAMKCTRESINSHISED-----PYKSNLYSLVGFGLLHNSYGSKSLF 65
QY 68 CLHVPEMPVR---GPDVEAYCLRCCKEYESSVTIKVTIILYLSILGLLLYVYLT 124
DB 66 CNPFGDSMPFRYTVSLDIKTRFL-CSCS-----GULLLFMDYLCV 104
QY 125 VEPILKR-RLFGLHAQLI 140
DB 105 ANPLTKRYRFLDHKSI 121

RESULT 6
AF0207
Probable membrane protein YP01702 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AF0207
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11598360
A:Accession: AF0207
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-450 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA90522.1; PID:gl5975733; GSPDB:GN00175
C:Genetics:
C:Superfamily: hypothetical protein H1671

Query Match 7.6%; Score 78.5; DB 2; Length 450;
Best Local Similarity 22.5%; Pred. No. 10;
Matches 23; Conservative 20; Mismatches 46; Indels 13; Gaps 4;

QY 38 EDVRCKICPPYKENSCHYKNIS-----QKDCCLHVEPMPVPGDVEAYCLRCECKY 93
DB 25 DDILCRATLYPMKI---HAIQRLSTARAQCCEC-DALFTPLPGNOTAYCPRCNAQI 80

QY 94 EERSSVTIKTIITIIYLSILGLLLYMYVLTLPVILKRLFG 135
DB 81 TSGRDWS-----LTRLTAMACMLLLIPFAFTPLISIRLLG 117

RESULT 7
T01011
Hypothetical protein At2g39810 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T517.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 23-Mar-2001
C:Accession: T01011; G84821
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, November 1997
A:Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.
A:Reference number: 214162
A:Accession: T01011
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-841 <ROU>
A:Cross-references: EMBL:AC003000; NID:g2642152; PID:g2642163
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
cuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: G84821
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-841 <STO>
A:Cross-references: GB:AE002093; NID:g2642163; PIDN:AA887130.1; GSPDB:GN00139
C:Genetics:
A:Gene: T517.11; At2g39810
A:Map position: 2
A:Introns: 133/2; 200/3; 365/3; 529/3; 575/3; 788/2
C:Superfamily: Arabidopsis thaliana hypothetical protein At2g39810

Query Match 7.6%; Score 78.5; DB 2; Length 841;
Best Local Similarity 22.7%; Pred. No. 19;
Matches 51; Conservative 27; Mismatches 74; Indels 73; Gaps 11;

QY 1 MATLWGGLRLGLSL--SCLALSILLQAQLDAKKNFEDVRCKICPPYKENSCHYIN 58
DB 398 MVLRSGRDGVSELVSGEAVTALRVREGLSEAFYQ-----RTCLKVKENN-----L 449

QY 59 KNTSQK-----DCD-----CLHVVPEMPVR-----GPDVEAYCLRC--E 90
DB 450 KNGAVKXASDDLDIWSWTEWELVNEFCCLSTRNLVDRIELPWNPEDEKYLHRCCLLD 509

QY 91 CKYEERSVTIKVTIITIIYLSILGLLLYMYVLTLPVILKRLRFGAQLIQSDDIDGDHQ 150
DB 510 SATDPSSAVGSLVLYFLQRYIQAQYVDL-----RLQKTEAFVSDNQIGE-- 558

QY 151 PFANAHVDLARSRRANVLNKVEYAQQRWKLQVQEQKSVFDRHV 195
DB 559 -----EVMFMRSS-----QSHW-----RKELVDRAI 579

RESULT 8
S08167
Balbiani ring 3 protein - midge (Chironomus tentans)
C:Species: Chironomus tentans
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: S08167
R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.
J. Mol. Biol. 211, 331-349, 1990
A:Title: The balbiani ring 3 gene in Chironomus tentans has a diverged repetitive stru
A:Reference number: S08167; MUID:90172404; PMID:1689777
A:Accession: S08167
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1700 <PAU>
A:Cross-references: GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058
C:Genetics:
A:Gene: BR3
A:Map position: 4
C:Superfamily: unassigned Balbiani ring proteins

Query Match 7.6%; Score 78.5; DB 2; Length 1700;
Best Local Similarity 24.5%; Pred. No. 38;
Matches 23; Conservative 9; Mismatches 39; Indels 23; Gaps 2;

QY 35 KNFEDVRCKICPPYK-----ENSGHIYNNKISQKDCDCLHW 72
DB 898 KKFNDFTSCGCPSPKLDCTGNKWSAETCTCGGVNRCGNLKNFNDNLQCCECKNQ 957

QY 73 EPMVPGDVEAY-CLRCEKYERSSVTIKVTI 105
DB 958 EMANCKSPRTWNYDTCKVCKNADSDDCVKPQI 991

RESULT 9
S73465
MG011 homolog D12_orf285 - Mycoplasma pneumoniae (strain ATCC 29342)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 10-Dec-1999
C:Accession: S73465
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia.
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73465
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-285 <HIM>
A:Cross-references: EMBL:AE000016; GB:U00089; NID:gi673796; PIDN:AA895787.1; PID:gi673
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: ribosomal protein S6 modification protein rimK

Query Match 7.6%; Score 78; DB 1; Length 285;

Best Local Similarity 28.6%; Pred. No.7.2;
Matches 22; Conservative 12; Mismatches 23; Indels 20; Gaps 3;

QY 124 LVEPIKRLFGHAQLIQSDDDIGHQPFPANAHVLAQR-----SRANVLNKV----- 172
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 30 LVDELNKKLVGHILL--DDETADHKHKNVELLIINRSRRIDFLTKHNFLNSFLINPQN 87

QY 173 -----EYAQQRWKLQ 182
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Db 88 IIVLVANDKIETVRWLXQ 104

RESULT 10
S75818
hypothetical protein slr1276 - Synecchocystis sp. (strain PCC 6803)
C:Species: Synecchocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #ext_change 20-Jun-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Ohta,
C., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
s.

A:Reference number: S74322; UID:97061201; PMID:8905231
A:Accession: S75818
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-275 <KAN>
A:Cross-references: EMBL:D90913; GB:AB001339; MID:g1653348; PIDN:BAAI8277.1; PID:g1653348
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: Synecchocystis hypothetical protein slr1276

Query Match 7.5%; Score 77.5; DB 2; Length 275;
Best Local Similarity 29.7%; Pred. No.7.6;
Matches 38; Conservative 14; Mismatches 41; Indels 35; Gaps 5;

QY 94 EERSSVTIKVTII-----YTSLGLLLLMVYLTLVEPIL 129
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 6 EYSSTT---TIIDEFDAGNYPTAFGITFTPOVGISLGILGALLANGYVLLNFVPAL 62

QY 130 KRRLFghAQLIQSDDDIGHQPFPANAHVLAQRSRVNLNVKEYYAQRWKLOVQEGRKS 189
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 63 -----GYQGIRQDEAKQAQVDQSQSEVLMRRLANAEV--QLGEAQR-KATVLELYAN 114

QY 190 VFDRHVVL 197
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 115 SEDLNTIL 122

RESULT 11
T42977
large tegument protein - ateline herpesvirus 3 (strain 73)
C:Species: ateline herpesvirus 3
A:Variety: strain 73
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T42977
R:Albrecht, J.C.; Fleckenstein, B.
submitted to the EMBL Data Library August 1998
A>Description: Primary structure of the herpesvirus ateles genome.
A:Reference number: 222274
A:Accession: T42977
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2471 <ALB>
A:Cross-references: EMBL:AF083424; PIDN:AAC95588.1
A:Experimental source: strain 73

Query Match 7.4%; Score 76.5; DB 2; Length 2471;
Best Local Similarity 22.4%; Pred. No. 85;
Matches 38; Conservative 28; Mismatches 81; Indels 23; Gaps 6;

27 LAOI-SDAKNFEDVRC----KCICPPKYENSGHYNNKINSOKDCDLHHVPEWPRGVDFE 83

QY 76 -----PVRG-----PDVEAYCLRCCKYERSVTIKVITIIYLSILGLL 115
Db 72 STBIALSVGRKRSFFVKGGTFIQRDIAAVKAVDDISFDLRKGETLG---LVGESCGKKS 128
QY 116 LLYVMVLTVEPIIKRLRF-----GHALIQSDDDIGDHQPFANADV 158
Db 129 TLGRCILRLPSPDQGVLFNGNDWTKLNARDMAARKHLQFVFQD-----PFASLHR 181
QY 159 LARSRSRANVLNKVEYAQQRWKQVQSQRSV 190
Db 182 MRKESIAEPLRISDLTQARKERVQEMRLV 213

RESULT 14

S46641

Probable membrane protein YJL193w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein J0349

C;Species: Saccharomyces cerevisiae

C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 19-Apr-2002

A;Accession: S46641; S56980; S56976

E;Purnelle, B.; Coster, F.; Goffeau, A.

Yeast 10, 1235-1249, 1994

A;Title: The sequence of a 36 kb segment on the left arm of yeast chromosome X identifies a gene ACO1 and two homologues to chromosome III genes.

A;Reference number: S46621; MUID:95274326; PMID:7754713

A;Accession: S46641

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-402 <PUR>

A;Cross-references: EMBL:X77688; NID:g1183982; PIDN:CAAS4767.1; PID:g547602

E;Purnelle, B.; Coster, F.; Goffeau, A.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56977

A;Accession: S56980

A;Molecule type: DNA

A;Residues: 1-402 <PUW>

A;Cross-references: EMBL:Z49468; NID:g1008405; PID:g1008406; MIPS:YJL193w

E;Obermaier, B.; Piravandi, E.; Rinke, M.; Domdey, H.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56937

A;Accession: S56976

A;Molecule type: DNA

A;Residues: 120-402 <OBE>

A;Cross-references: EMBL:Z49468; MIPS:YJL193w

C;Genetics:

A;Cross-references: SGD:S0003729

A;Map position: 10L

C;Keywords: transmembrane protein

Query Match 7.3%; Score 75.5; DB 2; Length 402;
Best Local Similarity 21.2%; Pred. No.18;
Matches 34; Conservative 22; Mismatches 47; Indels 57; Gaps 5;

QY 9 LRLGSLLSGLSIALSVLLLAQLDAAKNFEDVRCKICPPYKENSCH----- 55
Db 166 ITLILFSLCTLVLGWIIVQ-----EDNRGPASSNELREFSKYGVICAMLSMEIFV 217
QY 56 ---INXNI-----SQKDCCLHVVEPVRGPDVEAYCLRCCKYERSSVTI-- 101
Db 218 LQNIYKTVFTVRSQTSQNSGFSRQESPLPL-----YEKLDEKLVAK 262
QY 102 -----KVITIIYLSILGLLILYVMVLTVEPIIKRLRF 134
Db 263 KFKPSYDKLTLMIVISLVGFCLSFGWFITLFPVLFRYFF 302

RESULT 15

S48160

Metalloproteinase (EC 3.4.24.-) H-I precursor - carpet viper

N;Contains: disintegrin

C;Species: Echis pyramidum leakeyi

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000

C;Accession: S48160
E;Paine, M.J.I.; Moura-da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M.
Eur. J. Biochem. 224, 483-488, 1994
A;Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramidum leakeyi)
A;Reference number: S48160; MUID:95010025; PMID:7925363

A;Accession: S48160

A;Molecule type: mRNA

A;Residues: 1-617 <PAI>

A;Cross-references: GB:X78970; NID:g763092; PIDN:CAAS5565.1; PID:g763093

C;Superfamily: mouse meltrin alpha; disintegrin homology

C;Keywords: hydrolase; metalloproteinase; venom

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-192/Domain: propeptide #status predicted <PRO>

F;193-617/Product: metalloproteinase H-I #status predicted <MAT>

F;403-485/Domain: disintegrin homology <DIS>

F;337/Active site: Glu #status predicted

Query Match 7.3%; Score 75.5; DB 2; Length 617;
Best Local Similarity 34.0%; Pred. No.27;

Matches 17; Conservative 6; Mismatches 14; Indels 13; Gaps 2;

QY 46 CP-----PYKENSQHIYKNISQKDCCLHVVEPVRG-----PDV 82

Db 482 CELDVFQNGQPYQSNNGYCYNGCPILKNQCIHLWKFPAPGAVNVADPV 531

Search completed: March 15, 2004, 08:27:47

Job time : 23 secs